

NOTE TO USERS

This reproduction is the best copy available.

UMI[®]

A Neural Network Model of the Primary Visual Cortex

Alan Spara

B.Sc. Simon Fraser University, 2001

Thesis submitted in Partial Fulfillment of

The Requirements for the Degree of

Master of Science

In

Mathematical, Computer and Physical Sciences

(Computer Science)

University of Northern British Columbia

July, 2007

© Alan Spara, 2007



Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-48826-3
Our file Notre référence
ISBN: 978-0-494-48826-3

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

Many problems in modern computing require a visual component. That is to say, it is fairly common for applications to have a need to see their environments. These applications will typically employ techniques designed specifically to solve the particular task needed for the application, and have little or no relation to the human visual system.

Humans generally do not have difficulty interpreting the world around us. When traveling through known environments, we can easily recognize particular walls, doors and other objects in our view. We are not confused by the huge number factors that can complicate an image. The generalization and robustness of the human system would provide a huge benefit to any system that requires more advanced vision than is capable with the ad-hoc methods developed previously. If the underlying principles that make the human visual system so powerful can be identified and implemented programmatically, then a machine could reap the benefits obtained by humans.

The purpose of this thesis is to demonstrate that a visual system modeled after the human visual system will be robust and accurate enough to solve real world problems – and to be useful in a non-trivial application. By developing neural networks that directly model the most primitive image processing cells of the human visual system, a platform can be built on which advanced vision systems can be developed.

Table of Contents

Abstract.....	2
1. History of the Project.....	1
1.1. A New Vision System for Robot Navigation.....	1
1.2. Issues to be Resolved.....	3
1.3. The Path to Resolution.....	4
1.4. Comparisons to Other Research.....	5
2. A Model of the Human Visual System.....	7
2.1. Image Capture and Early Processing.....	8
2.2. Early Visual Processing.....	11
2.3. Visual Pathways.....	15
3. A Neural Model of the Early Visual System.....	16
3.1. Early Image Processing.....	17
4. A Model of Primary Visual Cortex	19
4.1. Cell Characteristics.....	20
4.1.1. Cell Configuration Summary.....	22
4.2. Cortical Maps.....	23
4.3. Complex Cells.....	24
4.4. End-Stopped Cells.....	24
5. Training Complex Cells.....	25
5.1. Training Cycle 01.....	26
5.2. Training Cycle 02.....	28
5.3. Training Cycle 03.....	29
5.4. Training Cycle 04.....	31
6. Training End-Stopped Cells.....	34
6.1. Training Cycle 01.....	35
6.2. Training Cycle 02.....	37
6.2.1. New Cell Configuration Summary.....	40
6.3. Training Cycle 03.....	40
6.4. Training Cycle 04.....	42
7. Off Angle Edges.....	44
7.1. Arc Region Transition.....	44
7.2. Issue Analysis.....	46
8. Result Analysis.....	49
8.1. Correct Line Interpretation.....	50
8.2. Correct Joint Interpretation.....	51
8.3. Curve Interpretation.....	53
8.4. Off-Angle Orientation Interpretation.....	54
8.5. Rounded Edges.....	57
8.6. End Stopped Cells and Joint Identification.....	59
9. Future Work.....	62
10. Appendix.....	64
10.1. Exact Target Angles.....	64
10.1.1. -75 Degree Results	64
10.1.2. -60 Degree Results	65
10.1.3. -45 Degree Results	65
10.1.4. -30 Degree Results	65
10.1.5. -15 Degree Results	65
10.1.6. 0 Degree Results.....	66
10.1.7. 15 Degree Results	66
10.1.8. 30 Degree Results.....	66
10.1.9. 45 Degree Results.....	66
10.1.10. 60 Degree Results.....	67
10.1.11. 75 Degree Results.....	67
10.1.12. 90 Degree Results.....	68
10.2. 20 Degree Line.....	69

10.2.1. -75 Degree Results.....	69
10.2.2. -60 Degree Results.....	69
10.2.3. -45 Degree Results.....	70
10.2.4. -30 Degree Results.....	70
10.2.5. -15 Degree Results.....	70
10.2.6. 0 Degree Results.....	71
10.2.7. 15 Degree Results.....	71
10.2.8. 30 Degree Results.....	71
10.2.9. 45 Degrees Results.....	72
10.2.10. 60 Degree Results.....	72
10.2.11. 75 Degree Results.....	72
10.2.12. 90 Degree Results.....	73
10.3. 23 Degree Line.....	74
10.3.1. -75 Degree Results.....	74
10.3.2. -60 Degree Results.....	74
10.3.3. -45 Degree Results.....	75
10.3.4. -30 Degree Results.....	75
10.3.5. -15 Degree Results.....	75
10.3.6. 0 Degree Results.....	76
10.3.7. 15 Degree Results.....	76
10.3.8. 30 Degree Results.....	76
10.3.9. 45 Degree Results.....	77
10.3.10. 60 Degree Results.....	77
10.3.11. 75 Degree Results.....	77
10.3.12. 90 Degree Results.....	78
10.4. Hallway Results.....	79
10.4.1. -75 Degree Results	81
10.4.2. -60 Degree Results.....	84
10.4.3. -45 Degree Results.....	87
10.4.4. -30 Degree Results.....	90
10.4.5. -15 Degree Results.....	93
10.4.6. 0 Degree Results.....	96
10.4.7. 15 Degree Results.....	99
10.4.8. 30 Degree Results.....	102
10.4.9. 45 Degree Results.....	105
10.4.10. 60 Degree Results.....	108
10.4.11. 75 Degree Results.....	111
10.4.12. 90 Degree Results.....	114
11. References.....	117
11.1. Papers.....	117
11.2. Human Spatial and Visual Systems Textbooks.....	122
11.3. Vision and Artificial Intelligence Textbooks.....	123

List Of Figures

• Figure 1 – A hallway image and an interpretation graph of that image	2
• Figure 2 – Early Dataflow of the Human Vision System	8
• Figure 3 – On and Off Response of Retinal Ganglion Cells	9
• Figure 4 – Vision Areas of the Occipital Lobe	11
• Figure 5 – Simple and Complex Cell Reactions	13
• Figure 6 – Layout of a Hypercolumn	14
• Figure 7 – A Model of the Human Vision System	17
• Figure 8 – Overlapping Receptive Fields	19
• Figure 9 – From Original Image to End-Stopped Cells	23
• Figure 10 – Complex Cell Positive Training Examples	25
• Figure 11 – Complex Cell Negative Training Examples	26
• Figure 12 – Testing Image	27
• Figure 13 – -50 Degree Cell and 10 Degree Cell Output	27
• Figure 14 – 10 Degree Target	28
• Figure 15 – Different 10 Degree Line Segments	28
• Figure 16 – Image Segments that Generate False Positives	29
• Figure 17 – Short Line Segments	31
• Figure 18 – Incorrect Processing at End Point	32
• Figure 19 – Positive End Stopped Training Examples	34
• Figure 20 – Negative End Stopped Training Examples	34
• Figure 21 – Original Dataflow	35
• Figure 22 – Modified Dataflow	37
• Figure 23 – Overlapping Positive Training Examples	39
• Figure 24 – Partially Activated End Stopped Inputs	40
• Figure 25 – Complex Cell Output with Differing Thickness	42
• Figure 26 – End Stopped Output with False Positives	42
• Figure 27 – New Negative Training Example	43
• Figure 28 – Off Angle Activation	44
• Figure 29 – Hard Separation between Neighboring Cells	45
• Figure 30 – Overlapping Cell Activation Regions	45
• Figure 31 – Shared Segments in Differently Oriented Cells	46
• Figure 32 – Actual Reaction of to Off-Angle Cells	47
• Figure 33 – Colours to Represent Complex and End Stopped Output	49
• Figure 34 – Correctly Interpreted Edges	50
• Figure 35 – Correctly Interpreted Joint	51
• Figure 36 – Activation Around a Curve	53
• Figure 37 – 23 Degree Edge Triggers Multiple Cells	54
• Figure 38 – 20 Degree Edge Partially Activates Cell	55
• Figure 39 – Partially Activated Cells to Off-Angle Edge	55
• Figure 40 – Edge Reconstruction Optical Illusion	56
• Figure 41 – Cell Reaction to Rounded Joint	57
• Figure 42 – Cell Reaction to Imperfect Edges and Joint	58
• Figure 43 – Location of End Stopped Cell Activation	59

1. History of the Project

1.1. *A New Vision System for Robot Navigation*

Robots that have been designed to navigate autonomously typically rely on Range Finders as their primary source of information about the world in which they are trying to navigate. Clearly, this is a very limited way of interpreting the robot's environment. The development of a vision system for these robots would give them a huge improvement to their navigation ability. In theory, an existing technique known as Markov Localization [Fox, Burard & Thrun, 1999] could be modified to use camera sensor readings rather than range finding images.

Markov Localization is a dense sensor approach to localization – meaning that it works by comparing the robot's current sensor reading to expected sensor readings. Expected sensor readings are pre-calculated for every possible position and orientation that the robot could have. The robot's position is then calculated by finding locations where the robot's current sensor reading matches the pre-calculated expected readings.

Dense sensor localization strategies traditionally rely on range finders as sensor readings. These sensor readings are very limited, but they do provide a number of benefits. First, the expected sensor readings can be easily predetermined from a map of the environment. Second, the actual sensor readings generated from the robot can be directly compared to these pre-calculated values. While a vision-based system would provide vastly more detailed information, neither of the advantages of a range finder apply anymore.

It is impossible to predetermine the expected sensor reading generated by a camera.

Since image data is so complex, the slightest change in circumstance can result in a vastly different image when doing a pixel by pixel comparison.

The key to making this process work is to develop a new vision system. This system will abstract away from the original image and build a graph that represents the image presented to the robot. Such a graph should be built in a manner that is independent of such conditions as lighting, robot orientation or slight changes in the area's configuration (such as a door being open or closed).

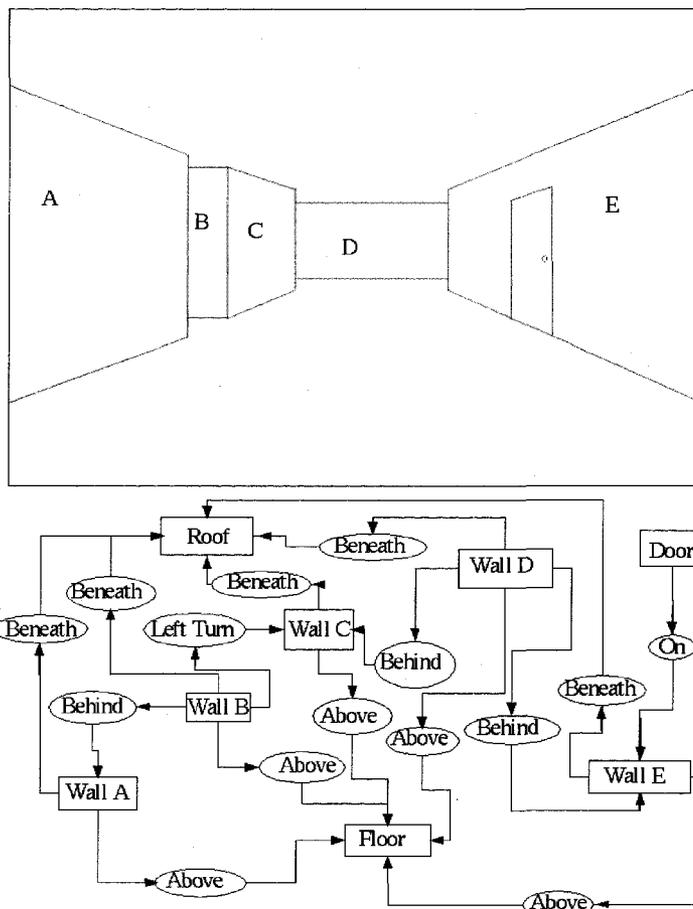


Figure 1 – A hallway image and an interpretation graph of that image. The original intent was to find a structure that could completely represent the original image in a manner that was insensitive to conditions such as lighting or camera position.

A well know structure that suits the needs of this work is the Conceptual Graph [Sowa, 1994]

Figure 1 illustrates a graph that might be used to illustrate a typical hallway image.

Every object on the original image is represented on the interpretation graph by a rectangular node. Oval nodes represent the relationships between those objects. From this graph, it is easy to see that wall B lies beneath the ceiling, above the floor, connects

to wall C with a left hand turn and disappears behind wall A. These sorts of relationships can be easily determined from a map of the environment and stored in a database of expected sensor results.

1.2. Issues to be Resolved

Clearly, the primary step in this theoretic vision system would be the ability for a robot to build a graph representing its current perceived image. In order to do this, boundaries between objects need to be interpreted in such a way as to make the objects identifiable. An edge map is well known technique for finding those boundaries. However, this edge map is simply a sketch of the original photographic image; there is no interpretation of that image. Building a graph representing the scene requires an interpretation of that edge map

It is known that an interpretation of an edge map can be built by a careful study of the corners in a scene. That is to say, the number and orientation of edges that come together to create a corner present a lot of information that can be used to interpret the objects defined by the edges. A joint that is comprised of three edges would most likely represent three surfaces coming together at a single point. These surfaces could include two walls and either the floor or the ceiling. By studying the angles of each of these edges, the exact configuration can be determined. The techniques for doing this are quite well known [Winston, 1992].

The underlying problems with using these techniques involve finding the locations of the joints, and identifying the types of those joints. In order to identify any particular joint, one must first identify the orientation of each joint from which the joint is built.

1.3. The Path to Resolution

As humans, we have no problems identifying objects in our environment. The boundaries between objects are quickly and easily identifiable. When trying to find methods to interpret those boundaries in an artificial environment, it follows that a reasonable starting point would be to examine those human systems.

A careful study of the human visual system reveals structures that seem tailor made to the problem at hand. It seems that the underlying building blocks of human vision are based on an ability to determine the orientation of boundaries between objects and to determine the start and end points of these boundaries (which will have a strong correspondence to the joints and corners in an image) [Mundel, Dimitrov & Cowan, 1997][Heitger, Rosenthaler et. al., 1992][Wurtz & Lourens, 1997] [Lowe, 2000][Dimitrov, 1998][Coren, Ward & Enns, 1996]. Some of the earliest portions of the human visual system consist of cells that react to boundaries of particular orientation.

Furthermore, other cells in this region react to the end points of the boundaries found by those directionally sensitive cells. These boundary determining cells form much of the underlying building blocks for our visual system. The primary theory behind the work presented in this thesis is that artificial neural networks can be build which model this functionality, and thereby produce a framework on which a robust vision system could be built for some artificial system (such as a robot).

A fully functional system that could build a graph representation of an input image is beyond the scope of a Master's Thesis. It would require modeling a huge variety of different parts of the human visual system. However, the fundamental building blocks of such a system would be a model of the early human visual system. This system is

responsible for segmenting the original image – identifying edges and edge ends (which correspond to potential joint locations) of the original image.

1.4. Comparisons to Other Research

Models of this portion of the human visual system have been attempted before [Heitger, Rosenthaler, von der Heydt et. al., 1992][Lowe, 1999][Wurtz & Lourens, 1997]. This research differs from those other projects in two principal ways. Other systems have modeled the primary visual cortex using mathematical operators, while my system functions through the use of neural networks. The other main difference is one of scale. Others have kept their models simple by limiting the number of different orientations recognized. In contrast, the neurons built here are designed to be a much closer match to the human system.

Other systems that model the underlying functionality of the primary visual system have done so with mathematical operations; functions that are applied to regions of a target image. This system will take a neural network approach to these models. This allows the research to abstract away from the mathematics, and instead concentrate on the underlying functionality. That is to say, this research can focus on the particular conditions that define a Complex or End-Stopped cell's activation, and not on the mathematics behind that definition. This will provide a better understanding why the cell activates rather than how the cell activates.

Another benefit of a neural design is that a neural network approach can be better scaled to accommodate future developments. For example, it is known that the orientation and

line ending cells of the human system rely on information fed-back from later stages of the vision system. The existing mathematical operators can never accommodate this feedback information. Instead, a completely new operator would have to be developed, completely invalidating the old model. However, this information can be incorporated into a neural approach. Once it has been determined that a neural approach to the human visual system is valid, those higher level visual systems can be developed. At that time, new neural networks can be built that use this information. While there are new networks, the underlying approach has not changed. Furthermore, as we learn more about how the human visual system that new understanding can be built into future neural networks.

The other main difference from other models of the primary visual system is that the cells developed in this research have orientations that are a much closer match to actual human cell activation. Determining how cells with neighboring orientations should react in relation to each other comprises much of the difficulty in developing this model. To put it another way, as an angle shifts from one orientation to the next, one must determine at exactly which point the first cell ceases to activate and exactly when the second cell begins to activate. When other researchers limit their models to only four or six different orientations, they reduce the complexity of this inter-cellular relationship.

2. A Model of the Human Visual System

The human visual system is a highly complex system that has evolved to be both highly robust and accurate. We are capable of easily identifying a wide variety of objects and placing those objects into their context in an extremely timely manner. We can recognize those objects even when they are occluded, have unexpected colour, are oriented in an unusual direction or are of a strange size. Furthermore, we are capable of doing this despite almost any particular lighting condition. In short, the human visual system provides a perfect example of what we would like an artificial system to be able to accomplish.

Researchers have collected details on the Human Visual System from several sources [Coren, Ward & Enns, 2004][Kolb & Whishaw, 1996][Rolls, Aggelopoulos & Zheng, 2003][Dimitrov, 1997]. For example, there is direct examination of the brain structures involved in similar mammalian vision systems. Inserting probes into a monkey's occipital lobe and studying the effects that cause a spike in electrical activity has provided considerable information. Less direct sources of information include case studies of people with damage to their visual systems. When a particular region of a person's brain is injured, there are often very specific results to that person's vision. Through these various sources, we have a fair understanding of how we are able to see the world around us.

The Human Visual System is composed of a large number of structures, each of which processes the image information in different ways. The image is passed from structure to structure; at each step being converted into a form which is closer and closer to

interpretation. A general overview of this process can be found in Sensation and Perception, 6th Ed[Coren, Ward& Enns, 2004].

The purpose of this section is to provide a basic explanation of some of these key structures; enough to build a model from which the system can be built. There are other structures which are not a part of this model (such as the Tectopulvinar Pathway). Details on these structures are not provided.

2.1. *Image Capture and Early Processing*

Any vision system must begin by capturing an image, and the human visual system is no exception. Once that

image has been captured, early portions of the human visual system process that image in order to make it easier to interpret.

Figure 2 illustrates a number of the structures in the human system, which will be detailed below.

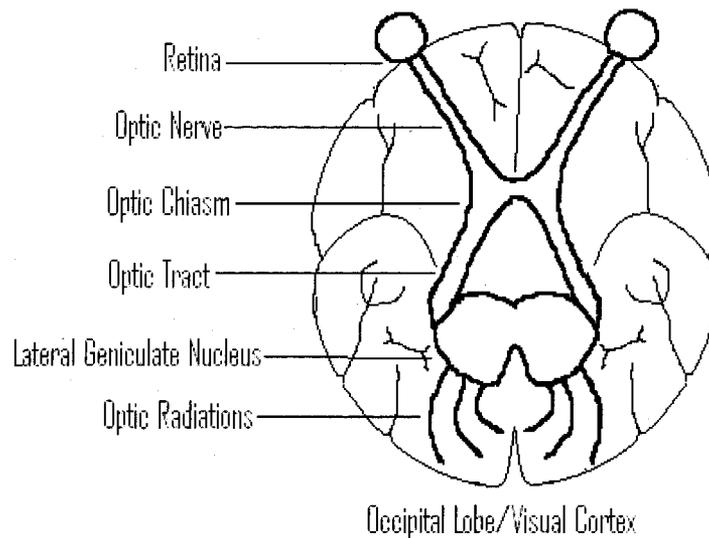


Figure 2 – Early Dataflow of the Human Vision System. Visual information is presented to the retina. It then undergoes several stages of intermediate processing as the image is passed from structure to structure until it reaches the occipital lobe, where the first stages of image interpretation begin.

The human system captures an image through a number of photo receptive cells located in the retina [Coren, Ward & Enns, 2004]. It is well known that we have two basic

types of photo receptive cells: rods and cones. These cells allow both nighttime and daytime vision. Rods activate in low light conditions, and do not allow for colour. In contrast, Cones react to daylight conditions. Furthermore, they come in three varieties, each reacting most strongly to different wavelengths of light. This provides our ability to see colour.

Almost immediately, there is some processing done on this image. The outputs of numerous nearby photo receptive cells are aggregated into various types of Retinal Ganglion cells, which transmit the image data to later stages of the vision system [Coren, Ward & Enns, 2004]. The set of photo receptive cells that cause a particular Retinal Ganglion cell to activate is known as the cell's Receptive Field. Neighboring Retinal Ganglion cells will have overlapping Receptive Fields; which naturally implies that any particular rod or cone will affect the output of many different Retinal Ganglion Cells. The original image captured by the retina is not maintained. Instead, the Retinal Ganglion Cells transmit to the Optic Chiasm this altered version; which represents an enhancement of the original image.

Different types of ganglion cells react differently to the intensity of light stimulus [Coren, Ward & Enns, 2004]. If a small point of light is moved through a cell's Receptive Field,

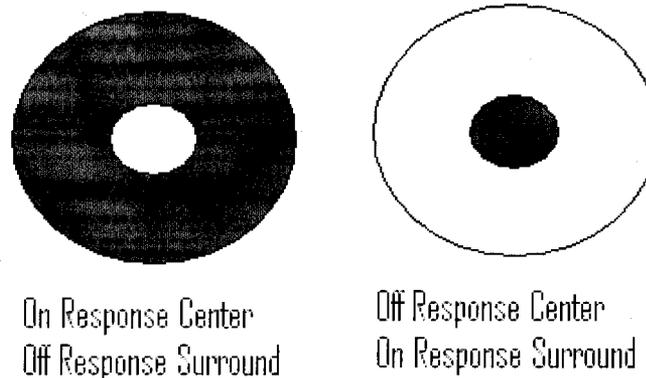


Figure 3 – On and Off Response of Retinal Ganglion Cells. Certain retinal ganglion cells produce a strong positive reaction to bright regions in the middle of their receptive field, and a strong negative reaction to dark areas outside of the center. Other ganglion cells produce exactly the opposite reaction.

the cell can either respond to the light being turned on (known as an On Response) or to the light being turned off (Off Response). Now, it seems as though some cells will produce an On Response when a point of light stimulates the center of the Receptive Field, and an Off Response in the outlying regions of the Receptive Field. Other cells will react in an opposite manner (off response to the center of the receptive field and on response to the outside). These two different types of response allow the visual system provide different processing of light and dark, and to better scale relative brightness and darkness of the image to make processing easier.

The ganglion cells (combined to form the optic nerve) transmit the image to the optic chiasm: a structure which separates the image into the left and right field of view [Coren, Ward & Enns, 2004]. These fields of view are then sent to be interpreted by the right and left lobes of the primary visual cortex respectively.

The Geniculostiate Pathway transmits this image to the Visual Cortex. The main structure of this system is the Lateral Geniculate Nucleus (LGN) [Coren, Ward & Enns, 2004]. The LGN integrates the processed image data coming from the Optic Chiasm with back projected image data from higher levels of the visual system, and transmits this to the various cortical maps of the Visual System through a series of cells known as the Optic Radiations.

2.2. Early Visual Processing

The Visual Cortex is located in the Occipital Lobe, at the very back of the brain. It consists of several sub-regions;

each responsible for a different type of analyses of the image

[Coren, Ward & Enns, 2004][Kolb & Wishaw, 1996]. For example, V3 is

used to process colour

information and V5 is used in processing motion. As is illustrated in Figure 4, these visual areas exist among the

folds and wrinkles in the Occipital Lobe at the back of the brain.

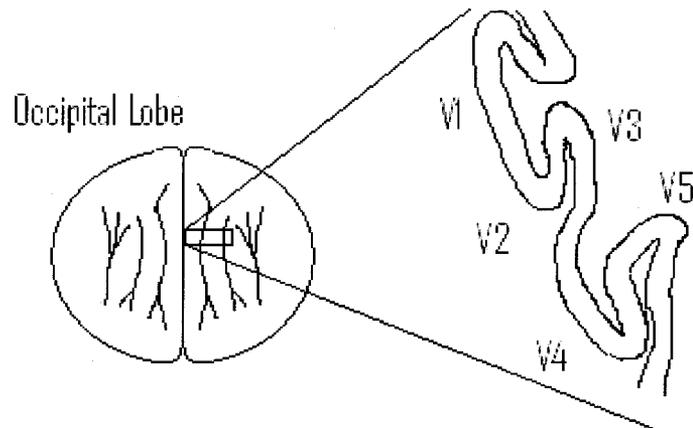


Figure 4 – Vision Areas of the Occipital Lobe. The now processed image is presented to several different cortical maps on the occipital lobe. These maps represent the first stage of image interpretation. Each region processes a different aspect of the image (such as intensity, texture, colour or motion). The Primary Visual Cortex combines the result of this processing to be used by higher level interpretation

Each of these sub-regions is a map of the original image presented to the retina. That is to say, there is a correlation between the location of photosensitive cells in the retina and the location of cells in each visual area which process that information. If exciting a particular cell in the retina causes some cell in a visual map to be stimulated, then one can estimate that a stimulating a nearby cell in the retina will cause a nearby cell in the visual area to be activated as well. Naturally, due to the processing of the image by the Retinal Ganglion Cells, this is not a direct 1-1 mapping, but a general relationship between the locations of the cells.

The primary visual cortex (known as V1) is located in the back of the occipital lobe. This area is the most important of the visual areas. As such, this is the area that will be the focus of this thesis. V1 is a highly complex region. It is organized into numerous layers that allow input from the LGN, input from other visual areas in the Visual Cortex and other higher processing centers of the vision system, and output to various neural streams.

The main purpose of the Primary Visual Cortex is image segmentation. It breaks down the original image into structures which define the boundary and joint conditions for that image. For example, when viewing a hallway image, V1 will produce reactions at the edges between walls or around doors or windows, and at the places where those edges meet. There are three cells used by V1 for image segmentation: simple cells, complex cells and end-stopped cells. These cells respond to differences in the intensity of the input image, but each in different ways.

Simple cells [Coren, Ward & Enns, 2004][Heitger, Rosenthaler et. al, 1992] respond to the boundaries between dark and light regions in the image. Simple cells are designed to react when one side of the cell's receptive field is dark and the other is bright. As such, these cells are somewhat reactive to different orientations. One simple cell will react to the transition from dark to light at a particular orientation, and the neighboring simple cell will react to the transition from dark to light at an orientation around 10 to 15 degrees away.

Complex cells [Coren, Ward & Enns, 2004][Heitger, Rosenthaler et. al, 1992][Mundel, Dimitrov & Cowan, 1997] also respond to oriented edges. These cells take the input from simple cells and perform further processing. There are two

types of Simple Cells which

could react to an edge,

depending on which side of

the edge is brighter and which

is darker. However, Complex

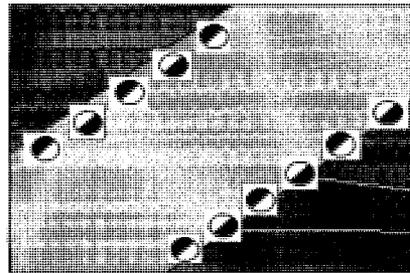
Cells will aggregate this

information, so only one

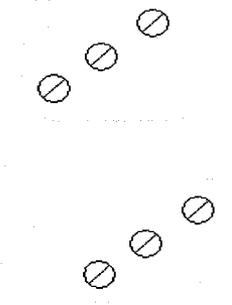
Complex Cell will be able to be

responsive to that edge. That

is to say, Complex Cells are purely orientation selective.



Simple Cells React to Original Image



Complex Cells React to Simple Cells

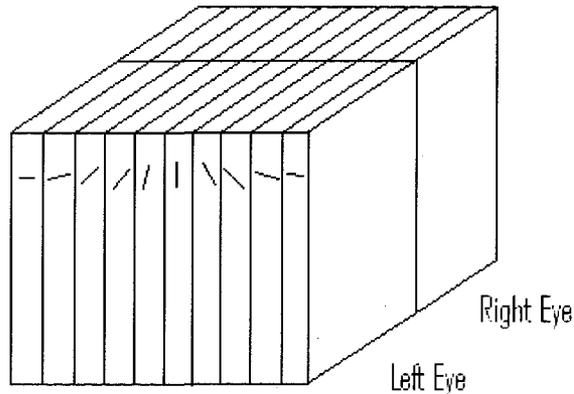
Figure 5 – Simple and Complex Cell Reactions. Simple Cells react to variations in intensity along a particular orientation. Complex Cells process the output of those cells to produce coherent depictions of an oriented edge.

End-stopped cells [Coren, Ward & Enns, 2004][Heitger, Rosenthaler et. al, 1992][Wurtz & Lourenz, 1997][Henricsson & Heitger, 1994][Mundel, Dimitrov & Cowan, 1997] are also directionally sensitive. However, they respond to areas where the edges end. End-Stopped cells have Complex cells as their input. The distinction is that End-Stopped Cells only react to the termination of a series activated of Complex Cells in the correct orientation. A short sequence of activated Complex Cells in an End Stopped Cell's Receptive Field will cause the End-Stopped cells to activate. However, a longer row of activated Complex Cells in that same field will have an inhibitory effect on the End-Stopped cell, and it will not activate. In this way, End-Stopped cells will activate at line terminations. It is also important to notice that they will often activate in the areas of the visual scene that contain edge joints – which contain

large amounts of information necessary to interpret the image. This makes these cells very useful as joint recognizers.

These cells are arranged in a very specific manner [Coren, Ward & Enns, 2004][Mundel, Dimitrov & Cowan, 1997]. There is one cell capable of responding at approximately every 10 to 15

degrees of orientation. These cells are grouped together and arranged in order. Successive cells in an orientation will react to successive orientations. This arrangement of cells is known as a hypercolumn.



A hypercolumn consists of cells which respond to a full 360 degrees of orientation.

Figure 6 – Layout of a Hypercolumn. The various orientation selective cells of V1 are arranged in a particular order. Cells of neighboring orientation are clustered in a sequence to form a group known as a hypercolumn. There a millions of these hypercolumns covering V1

Hypercolumns are scattered across V1, each responding to its own Receptive Field.

V1 also contains cells which operate across hypercolumns[Coren, Ward & Enns, 2004][Mundel, Dimitrov & Cowan, 1997]. These cells can suppress or enhance a detected edge based in information from surrounding hypercolumns.

Edge information from each different visual map is aggregated together to build a complete view of the edges and edge ends, which can then be passed on to the Object Recognition and Object Localization streams to produce a meaningful interpretation of the scene.

2.3. Visual Pathways

The object recognition portion of the human visual system is located in the temporal cortex, along the side of the brain[Coren, Ward & Enns, 2004][Li & Attick, 1994][Lowe, 2000][Rolls, Aggelopoulos & Zheng, 2003]. Cells have been identified in the inferior temporal cortex which respond to very specific features in the visual scene (such as a particular person's face). These cells are able to respond to the same feature from a variety of different positions and lighting conditions. In other words, they are insensitive to changes in size, orientation and other conditions

These cells have inputs that respond to less specific objects and shapes. For example, a specific face recognizer will respond to cells that react to general face characteristics.

These cells will react to cells which recognize simpler objects. This process continues in this manner back to cells that respond to very simple shapes – such lines and curves.

Ultimately, the pathway begins with the image parsing mechanisms of the early visual system.

The object localization portion is located along the top of the brain (and is therefore referred to as the dorsal pathway). It has connections to the object recognition stream, and to the internal map (stored in the hypothalamus). Like the object recognition stream, this stream begins with the segmentation information from the early visual processing,

3. A Neural Model of the Early Visual System

A complete vision system is beyond the scope of a Master's Thesis. Instead, the work will focus on the development of a working model of the Early Visual System. This model will lay the groundwork, from which future development can proceed. The image segmenting cells of the Primary Visual Cortex were chosen as the focus of this research because this is the first step in the Vision System which actually interprets the image. Earlier portions of the visual system alter the image to make it easier to be interpreted, but the image segmentation of V1 is the first step where any level of understanding of the image takes place.

If the image can be segmented in a manner similar to the segmentation done by the human system, a later process could be developed which would use the edges and edge terminators to find lines and joints. These can be used to build specific joint recognizers, which can be ultimately be used to build cells to recognize walls, doors, windows and other objects found in an image (modeling the successive levels of complexity found in the Human Object Recognition stream). These joints could then be combined into a comprehensive interpretation of the original visual scene (using a model the Object Localization stream of the human system). However, the first step must be to build the foundation for this processing – image segmentation.

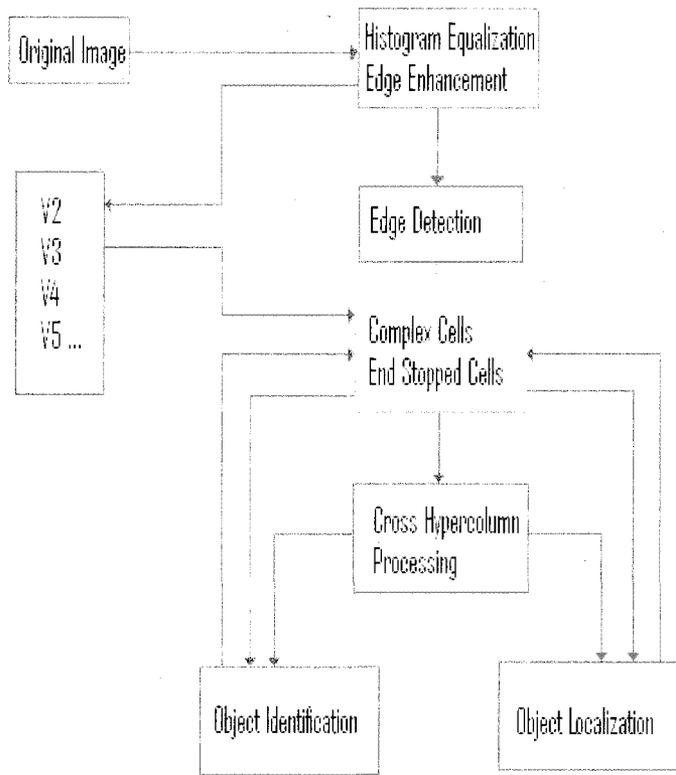


Figure 7 – A Model of the Human Vision System. In a manner similar to the actual human system, image data can be processed step by step in order to make the image easier to interpret. Neural Networks modeled after the Complex and End Stopped Cells can then be used to interpret the image. Further processing of the image is left to future work

The overall model used for building this system is detailed in Figure 7. The original image that is to be interpreted (a natural model for the photoreceptive cells of the retina) is processed with some basic techniques to simulate the early image processing of the Retinal Ganglion and other early cells; including the Simple Cells. The next step is the main

processing of the V1 model, the implementation of Orientation and Line Ending detectors. This leaves a lot of work that can be done in the future; such as the modeling of other early visual areas, cross-hypercolumn processing, higher level image recognition systems, and the incorporation of feedback from these systems back into the earlier processing steps.

3.1. Early Image Processing

The interpretation of an image with this model begins with some basic image processing steps. Since the current system is building a model of V1, there is no colour processing. Therefore, the image is first converted into black and white. A histogram equalization

routine is used to better accentuate the bright and dark regions of the image, and some basic edge enhancement is used to make the resulting image easier to interpret. These steps are done with basic, well known image processing techniques and do not require further explanation. This processing is similar to some of the early processing done by the Retinal Ganglion cells and other early vision systems.

The next part of this process is a somewhat controversial choice. Simple Cells have not been modeled with Neural Networks, as has been done with Complex and End Stopped Cells. Instead, it is done with a simple Edge Detection routine. The primary consequence of this decision is that the output map from this step has no orientation information, but a simple map that indicates where edges can be found for later interpretation.

There are several reasons for making this decision. It can be argued that the main point of the Simple Cell is edge detection, since those cells ultimately detect edges in a manner very similar to standard edge detection routines; by reacting to rapid changes in the underlying image intensity. It is also worth noting that other researches have followed this approach and have only modeled the Complex and End Stopped Cells.

However, the main reason for choosing to implement Simple Cells with a standard Edge Detection routine is that there is really no need for two sets of cells that react to different orientations; especially when processing a static image. Further processing routines do not require output from the Simple Cell. They rely only on the Complex and End-Stopped cells.

4. A Model of Primary Visual Cortex

The primary work in this thesis is the Primary Visual Cortex. That is to say, the directionally sensitive and line terminating sensitive cells that will produce the basis from which other vision systems can be developed. It has already been described how the early parts of the visual system will be modeled with some simple image processing techniques. The higher level processing streams will be left to future work.

What remains is the development of the Complex and End Stopped Cells that will model the cells found in the human Primary Visual Cortex. The full set of these cells will represent a logical hypercolumn, which can be used to interpret the edge map. These cells will be built using standard feed-forward neural networks.

The cells react to a Receptive Field, a small region which maps to a point on the original image. When a cell processes an image, it scans the entire Receptive Field, and marks the output of the cell at a matching position on the cell's output map. Every position on the input maps will be individually processed by the

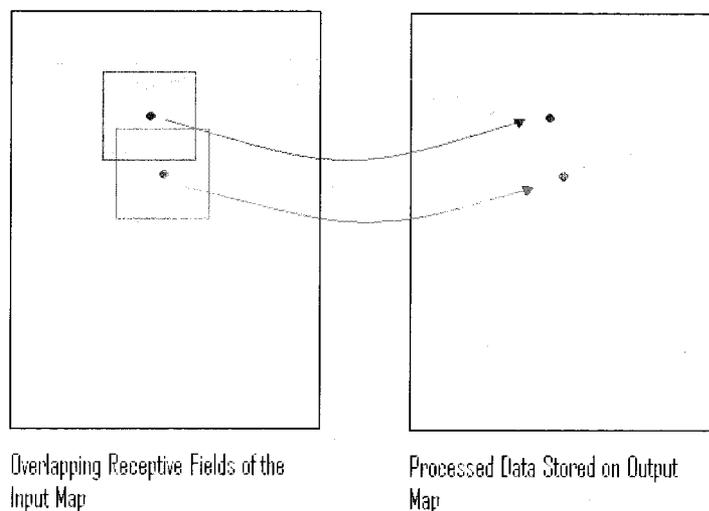


Figure 8 – Overlapping Receptive Fields. Every 9x9 region of the input is individually scanned with the Neural Network. The results of this processing is recorded on a direct pixel by pixel basis to an Output Map

hypercolumn, so the cell outputs will be the result of processing a large number of overlapping Receptive Fields.

It has already been mentioned that one of the benefits of neural network approach is that we can abstract away from the mathematics and focus on the conditions which will make the cells work correctly. This needs to be reiterated. Mathematics are tools which can be used to accomplish many interesting things. However, the maths used to create these applications should not be of primary concern, as they really are only a tool. Instead, we must focus on understanding the problem and each step taken to resolve that problem.

The mathematics behind neural network theory is well understood and there is no reason to reiterate them here. Instead, the focus will be on the conditions that will make the neural networks work correctly. Now, neural networks work through a training routine. They are given a series of both positive and negative training examples and are trained to respond correctly to those examples. If those examples have been chosen correctly, then the network will respond to previously unseen data in the correct manner. The work in this thesis will not focus on functions, formulas, calculations or algorithms. It will focus on the positive and negative training examples that will be used to create networks which respond correctly. It will also deal with incorrect behavior of those networks, and the modifications to the training examples that will improve that behavior.

4.1. Cell Characteristics

Before the artificial hypercolumn can be developed, the characteristics of the cells must be determined. There are two issues which must be addressed. First, the separation between the orientations which will stimulate neighboring cells (which will implicitly

determine the number of cells that will be built) must be determined. The other question which must be answered is the size of the Receptive Field, or the number of elements (pixels) of the input maps which will be processed by the neural networks.

The first decision that must be made is the number and orientation of the cells that will be used. The separation between orientation selective cells in the human visual system is between 10 and 15 degrees. One of the goals of this research is to build a V1 model which comes closer to the human system. So, the separation between the artificial cells is set at a similar level. In order to make everything work correctly, separation of 10 degrees is chosen, somewhat arbitrarily.

In both Complex and End-Stopped cells, one cell must be built to react at each possible orientation. However, the number of cells is different between the two types. Complex Cells are designed to react only to the orientation, and a line at any particular orientation is identical to another line which is 180 degrees away. As a consequence, only half a circle needs to be considered. As an example, there is no difference between a line with an orientation of 0 degrees and a line of 180 degrees, so only one of these angles needs to be considered. The model therefore calls for 18 Complex Cells.

In contrast, End-Stopped cells must consider angles from around the complete circle. This is because every line will have two end points; each pointing in opposite directions. Consider a line with orientation 30 degrees. This line will have two end points, one oriented at 30 degrees and the other oriented at -150 degrees. It follows that there must be twice as many End-Stopped cells as Complex cells.

In order to determine the size of the input region, several simple premises are considered. First, the cell should react to edges that pass through the exact center of the test region. To make this work, the region should be square and have odd number of pixels on each side. Secondly, the target model for each cell should be distinct for every different cell. Through some straightforward experimentation, it was determined that the smallest square region (with an odd number of pixels on each side) that could represent 10 degrees of separation is 9x9. Therefore, this is the size of the region used as input to the network.

4.1.1. Cell Configuration Summary

Receptive Field: 9x9 pixels

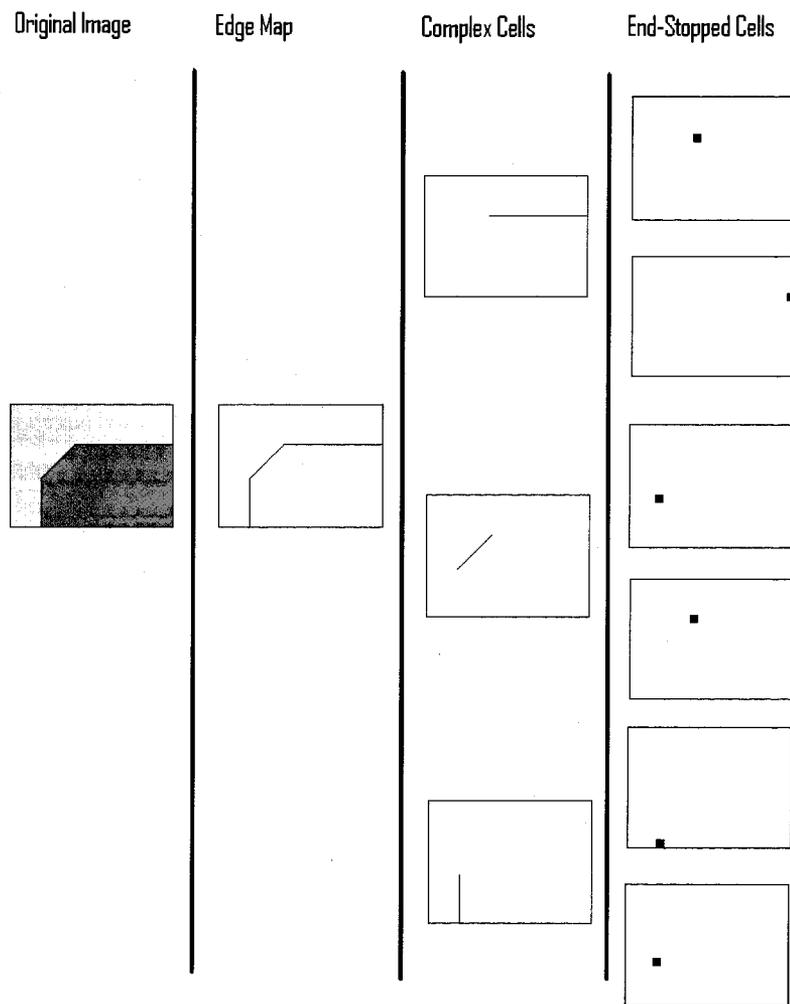
Orientation Separation: 10 Degrees

Number of Complex Cells: 18

Number of End Stopped Cells: 36

4.2. Cortical Maps

The output of the cells must be stored in a manner which is easy to interpret and is useful to later image processing steps. In humans, these cells are arranged in a cortical map. The most natural output for these cells is a series of images which represent a map of the cellular output.



If the region being processed contains an edge of the correct

Figure 9 – From Original Image to End-Stopped Cells. To interpret the original image, begin by generating an edge map. Complex Cells then isolate the sections of the edge map that lie in a particular orientation. Two End Stopped cells will then find the start and stop of that edge.

orientation, then it activates. The results of the scanning with these cells are stored on a map that corresponds to the original image. So, if the edge map is scanned at position (X,Y), then the results of that scan will be stored at position (X,Y) on the map that corresponds to that Complex Cell's output. There will be one output map for each Complex and End Stopped Cell.

4.3. Complex Cells

Complex cells will be designed to react to a small region of the edge map derived from the original image. The neural networks built will be designed to activate when an edge of the correct orientation is presented to the network. That is to say, when a line with the correct angle passes through the center of the network's receptive field, the cell activates. Any other image presented to the cell should cause it to fail to activate.

4.4. End-Stopped Cells

End-Stopped Cells work in a similar manner. They will scan a region the same size as the Complex Cells. However, they scan the results of the Complex Cells rather than the edge map. When a region that was previously determined to be of a particular orientation is found to terminate, then the End-Stopped Cell that reacts the orientation of the stopping edge reacts. In other words, the cell fires when the center pixel of the cell's activation region contains the last pixel of an edge with the correct orientation. The output from these cells is stored in another series of maps that correspond to both the Complex Cell output and the original image.

5. Training Complex Cells

Training examples are generated that will cause the cells to react correctly. The process used to create valid Complex Cells is cyclical in nature. Since the necessary training examples cannot be determined beforehand, a reasonable “first guess” at rules for generating training examples is created. Then, as any issue is uncovered, those rules will be altered to correct the problem. The networks created from these new rules are evaluated for any new issues, and a new set of training example rules is created.

Issues that may need to be addressed are not always obvious. In some cases, the output generated by the Complex Cell is clearly wrong and the training examples are in need of correction. However, in other cases, the issues with the Complex Cells cannot be recognized until those cells’ output is processed by the End Stopped Cells.

Figures 10 and 11 illustrate some examples of both positive and negative training examples. A network trained to respond to 30 degree edges will be trained to return a 1 for any of the images in Figure 10, and to 0 for any of the images in Figure 11.

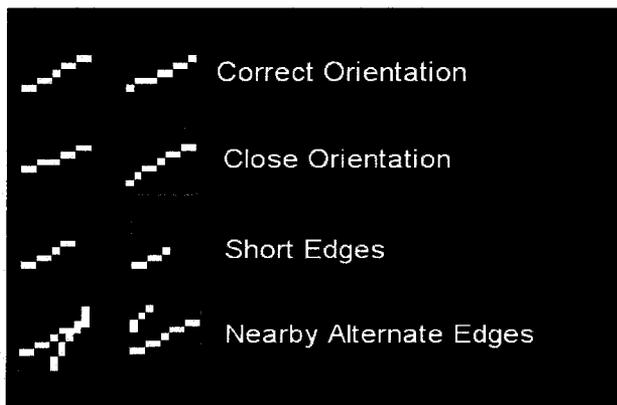


Figure 10 – Complex Cell Positive Training Examples. Some examples of small regions of an edge map that should generate positive Complex Cell output. These are presented to a network in order to teach it to respond correctly,

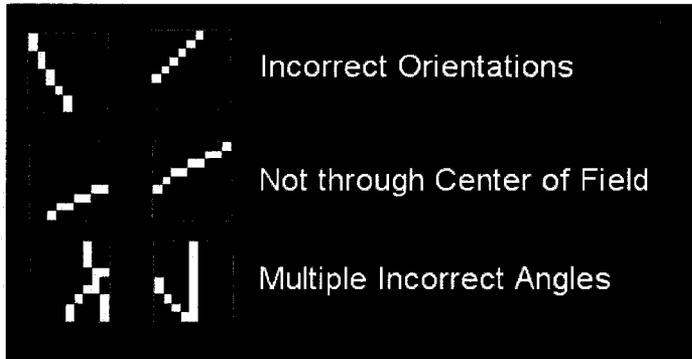


Figure 11 – Complex Cell Negative Training Examples. Some examples of small regions of an edge map that should not cause a Complex Cell to activate. These are presented to a network in order to teach it to respond correctly.

5.1. Training Cycle 01

The training examples used for the first version of the Complex Cells were designed using some basic rules that were chosen because they intuitively should produce cells that have somewhat correct results. A routine which takes a parameter and generated training examples was written. The training examples produced by this routine follow some basic rules.

- 1) A white line though the center pixel, oriented in the correct direction should cause the cell to react strongly (trains to 1)
- 2) The reversed image (white pixels turned to black, black pixels turned to white) always trains to not react (trains to 0)
- 3) A completely black region trains to 0
- 4) If the region contains only one or two random stray pixels, the cell should train to 0
- 5) Lines through the region, but orientated away from the specified angle should train to 0

- 6) Lines oriented in the correct orientation, but not passing through the center should train to 0
- 7) If there are two lines through the region, one correct and one with any other orientation and position should train to 1

Cells built based on these rules did indeed produce cells that were somewhat directionally sensitive. When presented with a test image consisting of several lines of various orientations, the line with the correct orientation is generally isolated. However, there are a couple of issues. First, there is quite a bit of noise around the joint. More significantly, the lines that have been isolated are not always complete – there are gaps in the line which causes a dashed-line appearance to the output.

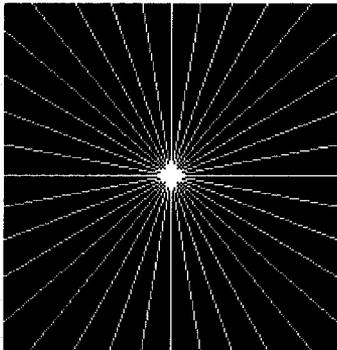


Figure 12 – Testing Image

Consider Figures 12 and 13. It is clear that the output for these cells do generally find the line with the orientation that they are designed to. However, the noisy joint and the dashed lines are particularly noticeable in the output for the cell oriented at 10 Degrees.

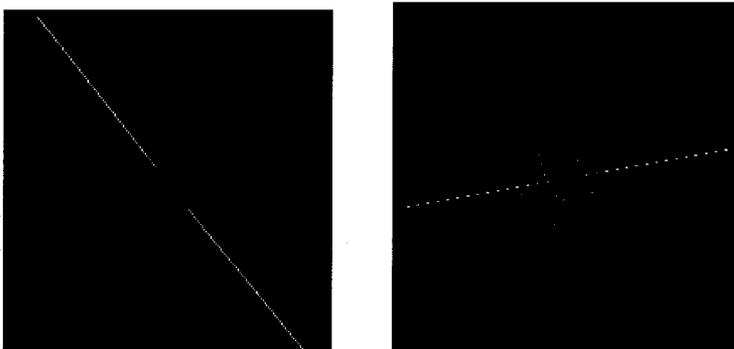


Figure 13 – -50 Degree Cell and 10 Degree Cell Output. The results of scanning the testing image with two different Complex Cells. Although the negative fifty degree cell reacts well, there are a number of problems with the ten degree cell's output.

5.2. Training Cycle 02

The dashed lines seen in the 10 Degree recogniser are an example of the first issue to be resolved. Although the 10 Degree line has been isolated, only certain portions of the line actually cause the cell to activate. As a result, the line is not correctly recognised. When the end-stopped cells are applied to this kind of output, they active all along the line rather than just at the ends.

This is a fairly straight forward problem, and is to be expected. While a line segment through a 9x9 region seems like a simple object to generate, it is in fact more complicated than one may initially expect. One must remember that a line is composed of pixels that are arranged in such a way as to be close to the line in question. Those pixels are, in many cases, not always aligned perfectly evenly. Consider, for example, figures 14 and 15. Figure 15 represents the line segment generated as the target for an edge with 10 degree orientation. That is to say, the 10 degree Complex cell is trained to respond with a 1 when it encounters a region of the edge map that matches this image.

However, as is illustrated by Figure 14, an actual edge with orientation 10 degrees is much more complicated. Different positions along the edge can look very different.

The clear resolution to this problem is to add

the missing positive targets to the training examples. While these different line segments

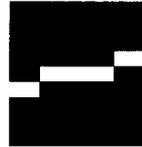


Figure 14 – 10 Degree Target

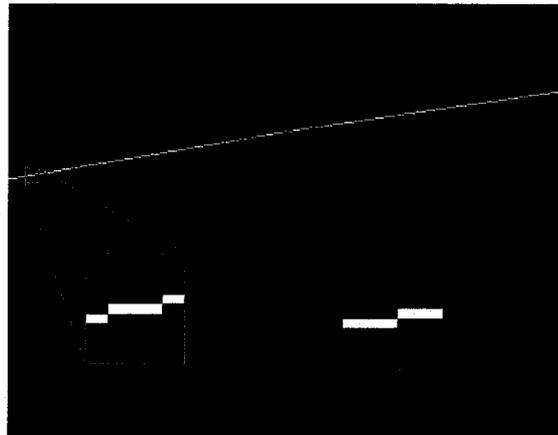


Figure 15 – Different 10 Degree Line Segments. Although both of these regions are a sub image of an edge with ten degree orientation, they look completely different from each other, and are different at almost every pixel.

may look very different, they all lie along the same orientation, and can be found along a much larger line in that orientation (see Figure 14). In order to correct this issue, a large line is drawn in the orientation of the desired cell. Then, a 9x9 region is scanned over this image. Every possible region that passes exactly through the center of the region is then added to the training examples with a target of 1. By adding these training examples to the routine that generates training examples, the dashed lines will be not be generated by the Complex Cells.

5.3. Training Cycle 03

The next step is to clean up the noise seen around the joints. This false activation was completely unexpected, as there is no apparent reason for the Complex Cells to trigger around the joint. In order to diagnose this issue, an analysis of the conditions that caused the cell to react must be taken.

First, a position where false activation takes place is found on the Cell Output map. Since a position on this map relates directly to a position on the original edge map, it is a straight forward process to find the image region that caused this false activation. This process is taken over several false activations of the cell. The edge map regions that caused the false activation are compared so that similarities can be determined.

Figure 16 displays some of the regions that are causing a Complex Cell to activate incorrectly. A cell that was trained to react to an edge of -50

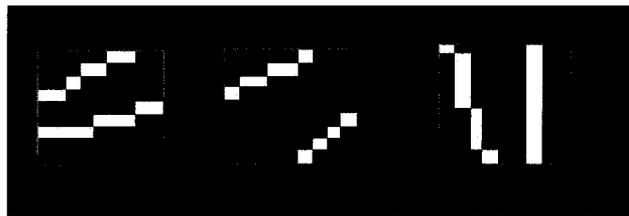


Figure 16 – Image Segments that Generate False Positives. It was not anticipated that two lines of completely incorrect orientation would generate a positive reaction, even when each individual incorrect edge would not

degrees was investigated, and it was found that these regions of the edge map were causing stray activation.

When looking directly at the edge map regions that cause false positive activation of the cells, it becomes immediately clear what is happening. While single edges that are off center and do not pass through the center of the region are specifically trained to cause the cell to not activate, this training does not generalize to multiple bad edges in the testing region. When two (or potentially more) completely incorrect edges fall into the region being tested, the cell can activate even though it would not activate if any of those lines were to be tested individually.

Specific training examples must be added to deal with multiple bad edges in the testing region. Combining multiple existing negative training examples into a single image did this. That is to say, every negatively oriented training example is considered, and paired with every other negative training example, one by one. Each of these pairs is then added to the set of negative training examples. Naturally, this increases the set of negative training examples by an order of magnitude (which slows down the training process considerably), but the results are worth the extra processing.

These extra training examples cause an unexpected side effect: the input space is no longer linearly separable. That is to say, it is no longer possible for a single cell to distinguish between correct activation and incorrect activation. Instead, Complex Cells must be modeled with a small network of interacting cells. This is likely a side-effect of the decision to model complex cells directly, rather than separate the functionality of simple cells. In the human visual system, the inherent (and somewhat ironic) non-

linearity of recognizing a straight line is processed through a multi-level orientation processing system, and it seems that a similar multi-level processing system must be built for an artificial vision system as well.

5.4. Training Cycle 04

Until this point, correcting issues with the directionally sensitive cells has consisted of re-training the cells until they match a straightforward understanding of how a cell should work. The next issue to be resolved does not have such a clear resolution, and the best solution can only be found through experimentation.

The cells need to have some mechanism for dealing with lines of a correct orientation, but do not span the entire testing region.

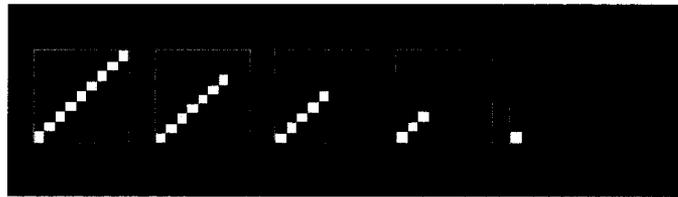


Figure 17 – Short Line Segments. Only edges that are long enough to reach the center of the testing region should be able to trigger a complex cell to activate.

Consider Figure 17. While it is immediately obvious that the first region displayed here must trigger a 45 Degree Complex Cell to activate, and that the last region should never cause that same cell to respond, it is not at all clear how the cell should react to the other three line segments.

There are several possible ways that the Complex Cells could be trained. First, they could be trained to only respond to a line segment that traverses the entire receptive field for the cell. That is to say, the first region would be trained to activate while the rest would be trained to not respond.

The concern with this approach has to do with the calculation of the End Stopped cell. When a long edge is processed by the Complex Cells, those regions of the edge which completely span the receptive field of the cells in that will activate those cells. However, the cells near the end of the edge will fail to activate, as the edge does not span the entire receptive field. When the End-Stop Cells process the output of these Complex Cells, it is clear that they will react at the wrong location, since the actual end of the edge was not recognized by the Complex Cells.

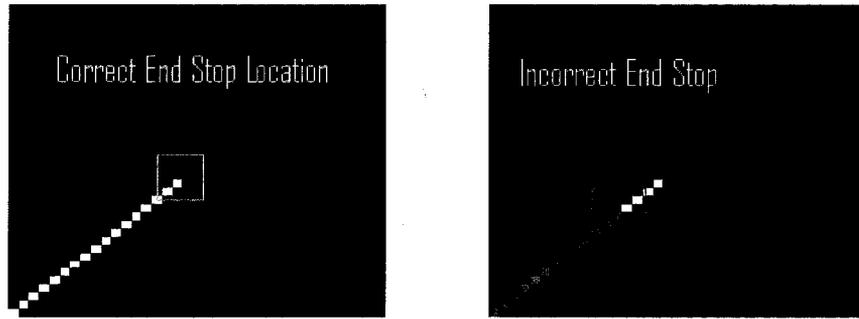


Figure 18 – Incorrect Processing at End Point. If only edges that span the entire receptive field is allowed to activate the Complex Cell, then the results will be shortened, and the End Stopped Cell will be triggered at the wrong location

The next option would be to allow the cells to provide partial reaction, scaled to react proportionally to the number of pixels left in the sample image. For example, the first test image of Figure 17 would provide 100% activation, the second with 78% (7/9 pixels in the image) and 56% activate for the third (5/9 pixels). However, it is unclear how the End Stopped Cells should react in this case; for similar reasons as the previous option. The line recognized by the Complex Cells will not terminate at the correct point, but will rather fade out over a longer distance. This will provide no clear point for the End Stop to react to.

The best results have been found when the Complex Cells react to any correctly oriented edge that passes through the exact center of the receptive field. Furthermore, all cells should be trained to respond with 100% activity. That is to say, a Complex Cell is reacting to an area around pixel (X,Y), then it should provide a full response whenever a

correctly oriented edge lies on the point (X,Y) ; even if it is only a partial edge. In that manner, the complete edge will be recognized by the Complex Cells and will provide good input to the End Stopped Cells that react in that area.

6. Training End-Stopped Cells

Training End Stopped Cells is similar to training Complex Cells. Rules are generated to produce training examples that can be used to cause the cells to react properly. Where Complex Cells were trained with simulated Edge Map values, the End Stopped Cell are trained with Simulated Complex Cell output. Once again, when it is found that the current set of rules do not produce correctly functioning End Stopped Cells, those rules are adjusted and a new set of cells is generated and tested.

Figures 19 and 20 display some examples of the training examples generated for an End Stopped Cell. These images represent a region of the output from the Complex Cell with matching orientation to the End Stopped Cell being trained

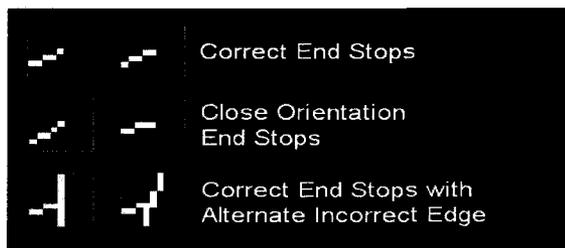


Figure 19 – Positive End Stopped Training Examples. Examples of cases where it is known that an End Stopped Cell should activate must be presented to the cell in order for it to be trained correctly.

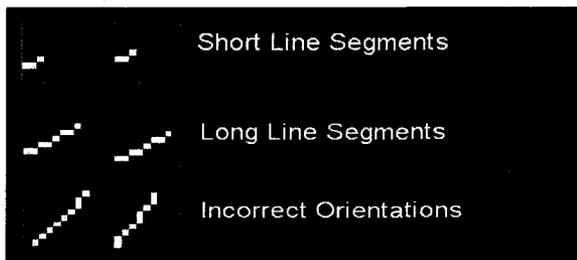
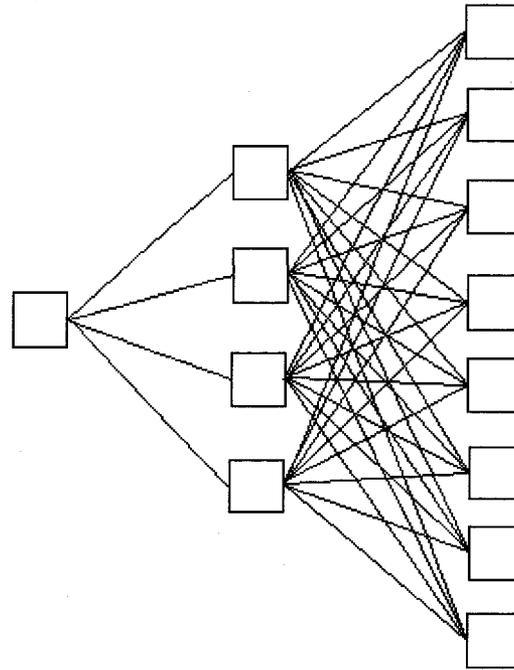


Figure 20 – Negative End Stopped Training Examples. Examples of cases where it is known that an End Stopped Cell should not activate must also be presented to the cell to make it work correctly

6.1. Training Cycle 01

The inputs of the first End Stopped Cells were designed strictly according to the model described earlier. That is to say, the outputs of all Complex Cells were aggregated into a single input vector and presented to the End Stopped Cell. This is illustrated in the following diagram.

It has already been pointed out that End Stopped Cells are directionally sensitive, as Complex Cells are. As a consequence, they must react strongly when the Complex Cell that corresponds to the End Stopped Cell's orientation displays an ending. The activations of the other Complex Cells have little impact.



Edge Map Complex Cell End Stopped Cell

The reality of this implementation is completely unfeasible. First, the size of the input vectors is quite large (size of the receptive field*number of Complex Cells).

Figure 21 – Original Dataflow. It is known that End Stopped Cells process the output of Complex Cells. However, it is not reasonable for an End Stopped Cell to process the output of every Complex Cell

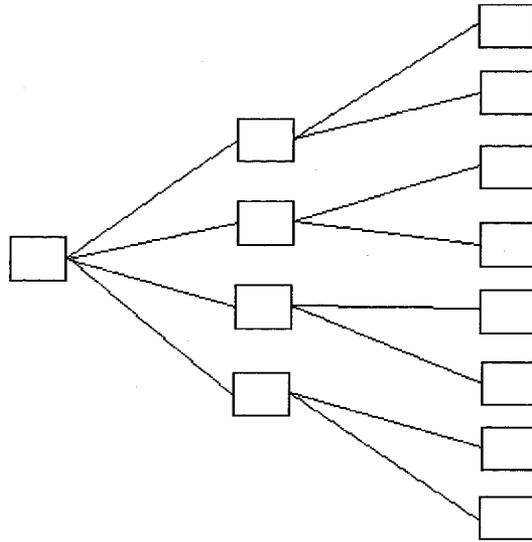
This dramatically slows down the speed at which the cell can react.

Even more importantly, the number of training examples needed to produce cells that react this way is completely unacceptable. First, the training examples for the matching

orientation Complex Cell must be calculated (according to rules found in the next section). Then, training examples must be generated for all of the off-angle Complex Cells, which will cause those inputs to be largely ignored. This is done with a large number of wildly varied input samples. When the cells are trained to this input, the off-angle cell inputs must not affect the activation of the cell, allowing the edge ending characteristics of the on-angle inputs to determine the activation.

The main problem with training End Stopped Cells according to this model is number of training examples that must be generated. Consider that every single on-angle training example must also train all other (off-angle) inputs to be irrelevant. This implies that there must be one training example generated for the combination of every on-angle training example with the combination of every off-angle example for every off-angle cell. This combination of combinations of training examples produces a total set that is so many orders of magnitude larger than just on on-angle training examples that it is not realistically possible to train cells to react in that manner.

Instead, a much simpler approach is taken. Since off-angle input cells should not affect the activation of the End Stopped Cell, they are completely removed from the input to that cell.



Instead, an End Stopped Cell will react only to the Complex Cell that matches its orientation. Since there are twice as many End Stopped Cells as Complex Cells, each Complex Cell will feed to

Edge Map Complex Cell End Stopped Cell

Figure 22 – Modified Dataflow. End Stopped Cells should only consider the output of the Complex Cell that matches its own orientation

two End Stopped Cells, but when processing an End Stop, only one Complex Cell needs to be considered.

Using this approach, the training can focus on the on-angle Complex Cells. The off-angle Complex Cells will be implicitly ignored simply by virtue of the fact that they are not taken into consideration.

6.2. Training Cycle 02

Generating training examples for End Stopped cells works in much the same way as the generation of Complex Cell training examples; a routine has been written to automatically produce both positive and negative training examples. Rules that specify both correct and incorrect exemplars are generated according to rules that will lead the trained cell to react accordingly.

The initial rules that are used to generate End Stopped training examples are listed below. All of these rules are used to generate training examples generated by Complex Cell with the matching orientation.

- 1) An edge of the proper angle that ends in exactly the center pixel should train to 1
- 2) White regions with black lines always train to 0
- 3) All black regions train to 0
- 4) One single activated pixel in the receptive field is just noise, and should train to 0
- 5) A line of the correct orientation, but not long enough to reach the center of the receptive field should train to 0
- 6) A line of the correct orientation, but too long (i.e., extending well past the center of the receptive field) should train to 0
- 7) Correctly oriented lines that miss the center of the receptive field should train to 0. This includes lines which are either too short, the correct length or too long.
- 8) Lines with the incorrect orientation train to 0. This includes edges that are in the center of the receptive field as well all cells that are offset from the center of the field.
- 9) A correct end stop should train to 1, even when there is also an incorrectly oriented line in the receptive field. That incorrect line could also be offset from the center of the receptive field
- 10) A line segment which has a very similar orientation to the correct one, and stops exactly at the center of the receptive field should train to 1

The first set of cells that used these rules could not be trained. Every time a cell was generated using these rules, a large number of the training examples would always be incorrectly mis-classified by the resulting networks.

To try and determine the cause of this network error, a careful study was made of the training examples

created with these rules. The problem arises from some

very basic training examples. Every

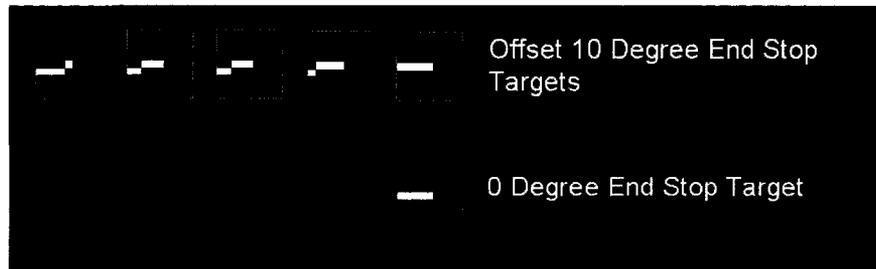


Figure 23 – Overlapping Positive Training Examples. Although a 9x9 region can distinguish ten degrees of separation for a Complex Cell, it is not sufficient to maintain separate input values for the End Stopped Cell. It is seen here that both zero and ten degree cells will react to the same input.

different line segment that could be a part of an edge with the correct orientation must cause the cell to activate if it ends exactly on the center of the receptive field. The problem is that, in some cases, there is no difference between a necessary training example for one cell and a necessary training example for its neighbor's cell (which must train to 0).

As can be seen, there are a large number of line segments that could trigger a 10 Degree End Stopped Cell. In fact, there is such a wide variety of examples that some are indistinguishable from necessary 0 Degree End Stopped Cell targets.

As was mentioned previously, the configuration of the cells (10 degrees of separation between orientations, with 9x9 pixel receptive fields) was chosen based on two criteria: the orientation is similar to the separation between cells in the human visual system and 9x9 was sufficient to separate these edges. However, this analysis was based on

Complex Cells, not on End Stopped Cells. It seems that the short line segments required by End Stopped Cells cannot be separated with this configuration.

In order to generate required cells, the configuration of the cells must be altered. There are two options for making this alteration: the orientation between cells can be increased or the size of the receptive fields can be increased. The decision on which option to take is a little bit arbitrary. Since 9x9 is already a fairly large receptive field, and there is some room in the cell orientation separation (the human cells tend to have a separation between 10 and 15 degrees of separation), the angle between cells was increased to 15 degrees.

6.2.1. New Cell Configuration Summary

Receptive Field: 9x9 pixels

Orientation Separation: 15 Degrees

Number of Complex Cells: 12

Number of End Stopped Cells: 24

6.3. Training Cycle 03

Until this point, all cells have been built assuming perfect input in the hope that the cells would generalize to cases where the input has some unexpected conditions. This works well for Complex Cells, since edge maps produce quite good input to the cells. However, Complex Cells do not always produce perfect response to End Stopped Cells.

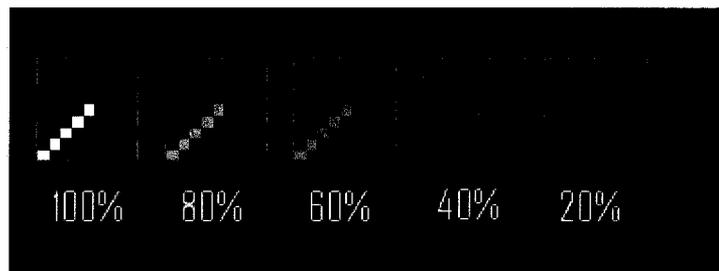


Figure 24 – Partially Activated End Stopped Inputs. When the Complex Cell's output is not at perfect strength, the End Stopped Cell's output should be scaled to match

When the image presents unexpected edges to the Complex Cells, occasionally they will produce partial activation. That is to say, rather than generating a 0 or a 1 the cell will react with some in between value, such as 0.6 or 0.7. When an End Stopped Cell is presented with this partially activated cell as input, they can produce unexpected results.

In order to deal with these partially activated cells, new training examples were generated to represent positive training examples which are partially activated. The target for these partially activated inputs is scaled to match the activation level of the cell input. So, if the input to the End Stopped Cell is only 80% of the normal input then the cell will react with 80% activation. If the Complex Cell only provides 40% input then the End Stopped Cell will provide 40% activation.

6.4. Training Cycle 04

The next issue to occur when building End Stopped Cells is another issue that is caused by unexpected processing of the Complex Cell. It seems as though, in certain circumstances, an off angle edge can produce some stray activation around the edge, which causes the Complex Cell output to alternate between somewhat thicker and thinner sections of the edge that it finds.

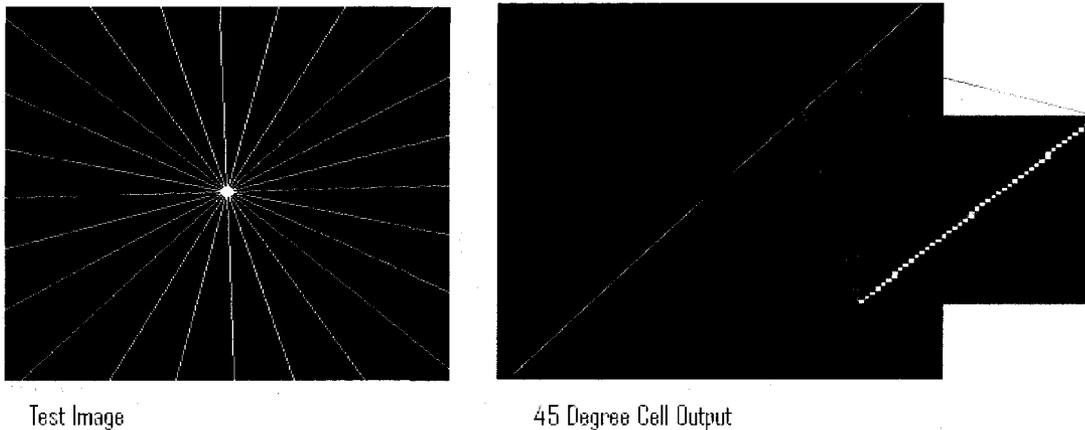


Figure 25 – Complex Cell Output with Differing Thickness. When an off angle edge is composed of small jagged lines, small imperfections in the Complex Cell output can arise.

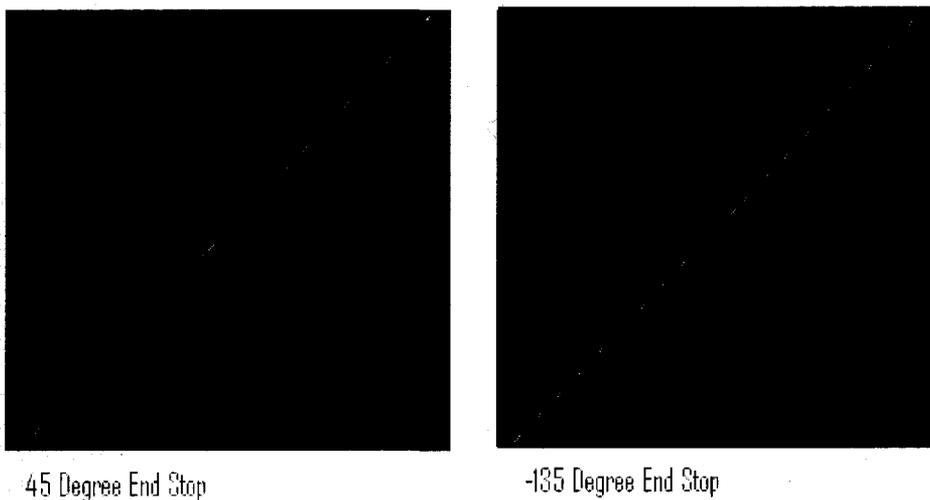


Figure 26 – End Stopped Output with False Positives. These imperfections in the Complex Cell can produce false activation in the End Stopped Cell, which can be trained out with some particular training examples.

It seems that the transition between the thicker and thinner sections of the Complex Cell output is sufficient to stimulate the End Stopped Cells, at least to a small degree.

There are two options that can be used to correct this issue. It is reasonable to believe that this is an issue with the Complex Cell. However, there are no straight forward techniques which can separate out the false positives in the Complex Cell Output from the correct cell output. Instead, this issue can be resolved by adding think-to-thin negative training examples to the End-Stopped Cell training examples.



**Figure 27 – New
Negative
Training
Example**

7. Off Angle Edges

Once the cells are working reasonably well, there is still quite a bit of work that must be done to fine tune them to ensure correct results. The main issue when creating these cells concerns edges that do not lie exactly along the orientation that the cell is designed to respond to. In any realistic situation, most edges will not line up exactly with the cell's primary orientation. They must be able to respond to angles that are nearby to the correct orientation. In essence, these cells need to react not to a particular output angle, but to a region of arc.

7.1. Arc Region Transition

Clearly, a cell should react to an edge that is very close the main target angle. Also, the cell should not react to an edge that is very close to the next cell over. However, it is not clear how the cell should react to an edge that lies immediately in between the two cells' main orientations.

Consider Figure 28. Orientation A and Orientation A+1 are adjacent to each other, and each has a cell trained to react to that orientation. Now, if cell A is presented with an edge that is some small orientation away from cell A's main orientation (in green), then

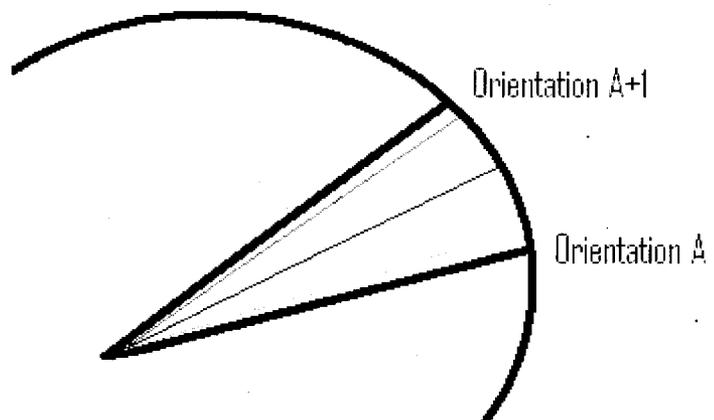


Figure 28 – Off Angle Activation. Most edges do not lie exactly on the main orientation of a cell. Instead, rules must be developed with determine how off angle edges transition from one cell to the next.

clearly cell A should react. However, if presented with an edge that is much closer to cell A+1 (in red) then cell A should not react. The issue in question is in regards to an edge (in blue) that lies immediately between cell A and cell A+1s' main orientation.

There are many possible ways of dealing with this situation. The most obvious approach is to try and draw a hard boundary between these two cells' activation regions. An angle is chosen that lies in between Orientation A and Orientation A+1, call it Orientation A+1/2. Now, if the edge being tested by cell A has an orientation less than Orientation A+1/2 then it activates. If that edge has an orientation greater than Orientation A+1/2 then it does not.

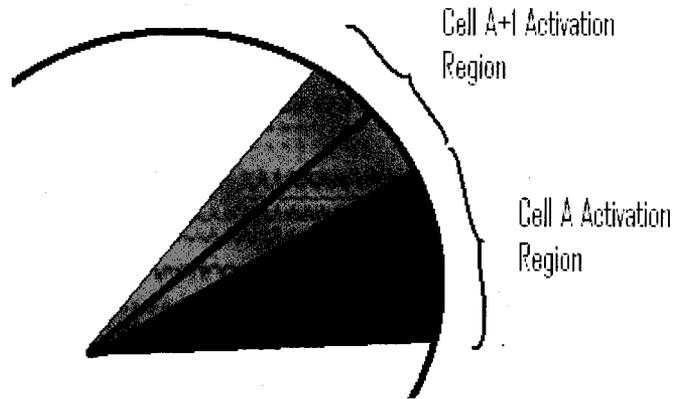


Figure 29 – Hard Separation between Neighboring Cells. Early cells attempted to classify an edge with a particular cell. A hard decision boundary was established between neighboring cells, and an edge was to be classified according to which side of the boundary it fell on. This approach was doomed to failure.

The other option is to have some overlap in the reactions between the cells. In this case the edge that lies right between the two cells' main orientation will cause both cells to react.

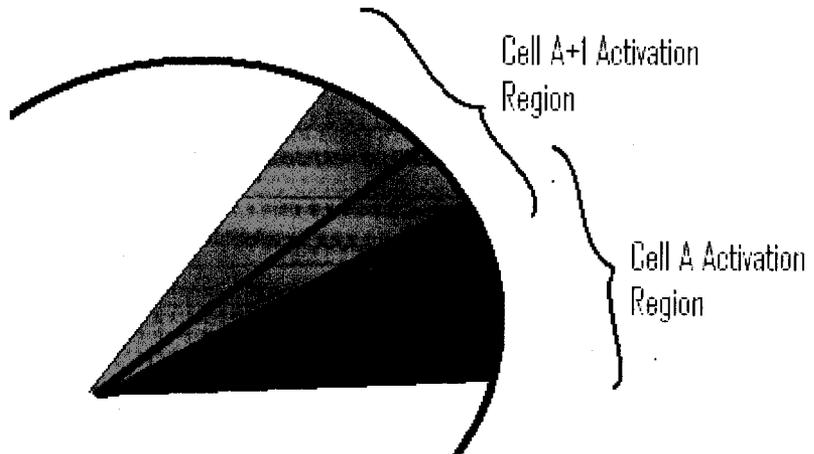


Figure 30 – Overlapping Cell Activation Regions. Later cells incorporated a Joint Activation Region – a small set of angles that could stimulate both neighboring cells.

This can be illustrated with Figure 30. Cell A is trained to react to the blue

region. Cell A+1 is trained to react to the red region. In addition, there is a purple

region. Both cells will react to an edge whose orientation lies in the purple region. From here on, this purple region will be known as the Joint Activation Region.

7.2. Issue Analysis

Through experimentation, it was found that the second option is the only one which can actually work. The problem with attempting to define an orientation that serves as a hard boundary between two cell's output regions is that small segments of differently oriented lines can look

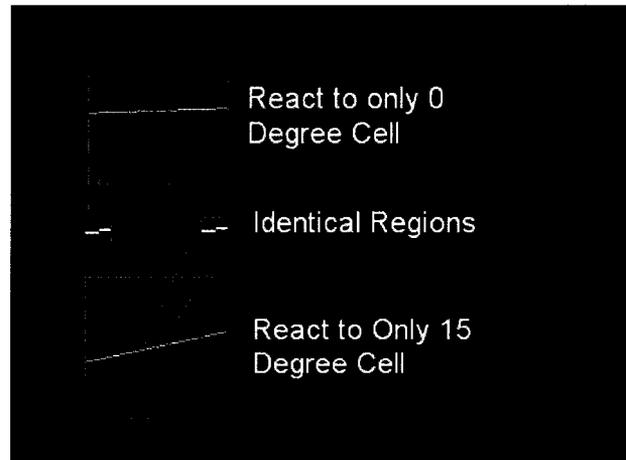
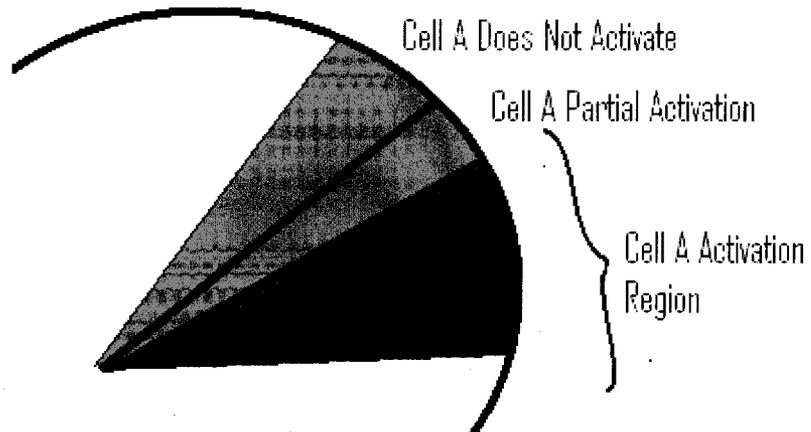


Figure 31 – Shared Segments in Differently Oriented Cells. Although these two edges should be classified by different cells, sub regions of these edges are identical.

identical. When two lines have similar but different orientations, it often happens that certain segments of those edges will look exactly alike. The consequence is that an edge on one side of the hard cutoff line will have areas that look exactly like areas on the other side.

Attempts to build a network that has a hard cutoff at a particular orientation can never succeed because they will be required to be both active and inactive for a particular line segment. Consider Figure 31. While the two edges have very different orientations, and should produce reactions to different cells, there are segments from those edges that are identical and cannot be separated.

In fact, there is no real method to separate the two regions in a perfectly clean manner. No matter where the border between the two regions lies, there will be regions immediately outside of the activation region which will have segments that still trigger the cell. Instead, the cells will react with degrading activation as the edge being tested is further away from the cell's target angle. It has been found experimentally that the best results are



obtained when there is a small area where two neighbor cells will both

Figure 32 – Actual Reaction of to Off-Angle Cells. In order to produce cells that operate properly, both Joint and Partial Activation regions must be established. Joint regions stimulate both neighboring cells. Partial activation cells may cause some stray activation of the wrong neighboring cell, but not enough to later trigger the End Stopped Cell to react.

react. Edges outside of the joint activation region provide partial reaction throughout the activation region for the neighbor cell. In this case, partial activation implies that certain segments of the edge will cause the cell to activate and other segments will not. This Partial Activation Region happens implicitly; there are no special training examples needed to cause it.

When training cells, the goal is to carefully balance the on and off response of the cell. This implies that training examples must be chosen in order to meet some very specific criteria

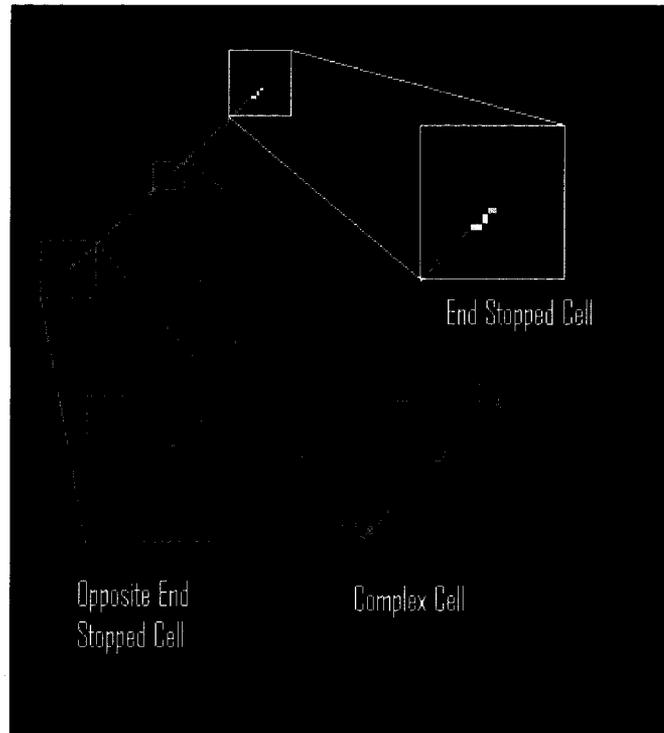
- on responses must be clearly separated from off responses
- a cell should never produce an on response for its neighbor's primary orientation

- the region of joint activation between neighboring cells must be minimized
- For complex cells, the vast majority of the partial activation region should provide only stray activation, which is insufficient to stimulate the End Stopped Cells.
- For End Stopped Cells, there should be a minimum amount of stray activation caused by the partial activation regions

Cells that meet these conditions have been found experimentally. For Complex Cells, a shared activation region of approximately 5% to 10% of the total activation region produces reasonably good results. End Stopped Cells must be built with a smaller joint activation region. If more than 2% to 5% of the total activation region is trained as a joint activation region it becomes inevitable that a cell will be trained to react to its neighbor's primary orientation.

8. Result Analysis

The following sections display images which have been included to demonstrate the functionality of these cells. They have been created in order to demonstrate how the different cells react to specific areas of the edge map. The results of the Neural Network outputs have been combined into a single image. The Complex Cell results have been loaded into the red component of the image. The green and blue



components contain the results for the pair of corresponding End-Stopped cells (for the angle that

matches the Complex cell and 180 degrees opposite). A point that only stimulates the Complex Cell will be marked in red. A point that stimulates the End-Stopped Cell of the same orientation as the Complex Cell should be marked in yellow (red + green). A point that stimulates the End-Stopped Cell 180 degrees opposite of the Complex Cell orientation should be marked in magenta (red + blue).

These images have been created specifically to demonstrate the activation of these cells in certain circumstances. They are heavily processed in order to demonstrate this functionality, but they do represent a true and complete record of the cell output in that area. The unaltered outputs to the testing situations have been provided in Appendix A.

Figure 33 – Colours to Represent Complex and End Stopped Output. Red indicates that a Complex Cell has been activated, yellow shows where an End Stopped Cell is active and magenta indicates that the oppositely oriented End Stopped Cells is active.

For the sake of brevity, throughout this analysis the Complex Cell and the two matching End Stopped Cells will be referred to only by the orientation that the Complex Cell has been trained to respond to. For example, the 30 degree cells will include the Complex Cell trained to 30 degrees, the End Stopped Cell trained to 30 degrees and the End Stopped Cell trained to -150 degrees.

8.1. Correct Line Interpretation

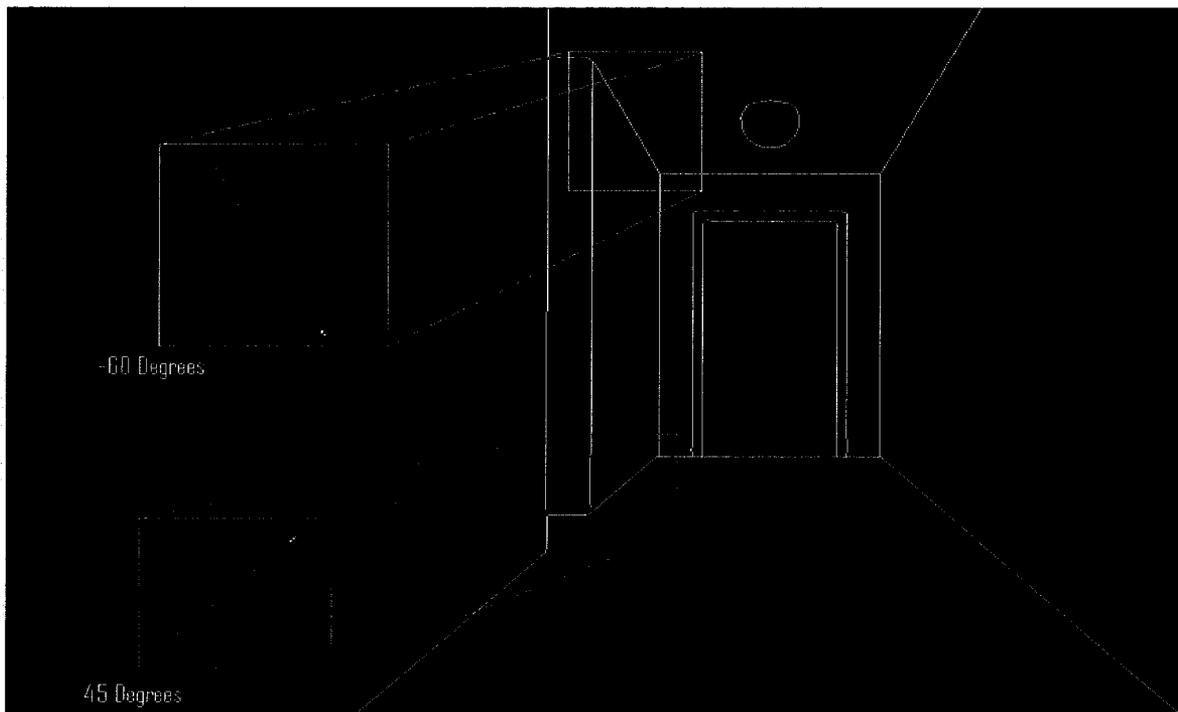


Figure 34 – Correctly Interpreted Edges. An edge is correctly classified by a connected series of Complex Cell outputs, bounded by both matching an oppositely oriented End Stopped Outputs

Figure 34 shows how the cells react to an edge. Proper interpretation of an edge should consist of a complete red line along the edge (Complex Cell Activation), with a yellow mark (Complex Cell and matching orientation End-Stopped Cell) at one end and a magenta mark (Complex Cell and opposite angle End Stop) at the other.

This is exactly what the output of the cells show. The 45 degree cells clearly separate out the boundary between the wall and the floor, while the -60 degree cells have found the wall/ceiling boundary. The ends of these boundaries (where the edges combine with other edges to form joints) have been clearly labeled in yellow and magenta.

8.2. Correct Joint Interpretation

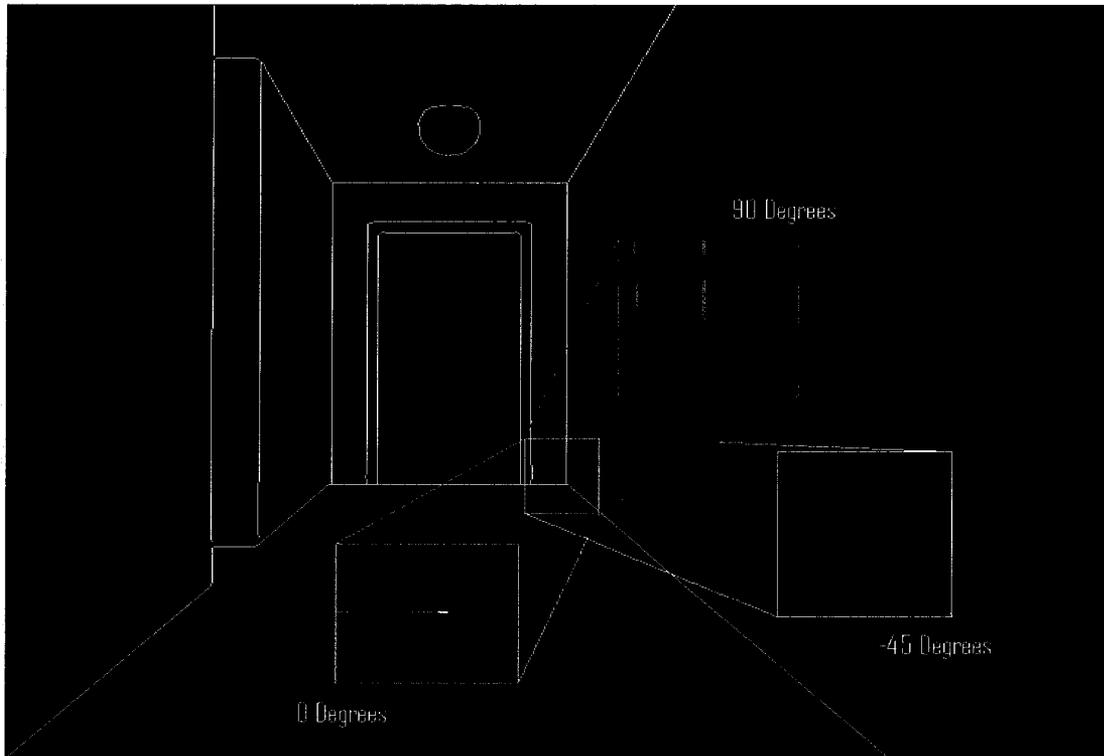


Figure 35 – Correctly Interpreted Joint. A joint is classified by a point where several differently oriented End Stopped Cell outputs are all triggered.

Figure 35 displays the correct interpretation of a joint. Since the joint is clearly defined in the underlying edge map, the joint is easily found.

The Complex Cells have clearly identified the edges of the correct orientation. Notice that each cell only finds the edge(s) that matches that cell's orientation.

Furthermore, there is End-Stopped activation correctly specified at the joint generated at the end of each line. For example, the output of the 0 degree Complex Cell has caused the 0 degree End-Stopped cell to activate. This can be seen by the yellow mark at the end of the 0 Degree output. In contrast, both the -45 and 90 degree Complex Cells have caused activation of the End-Stopped Cell that is 180 degrees opposite of the Complex Cell orientation (135 and -90 degrees respectively). These outputs can be clearly seen in the magenta spots at the end of the matching output.

A joint can be seen as a point where several End-Stopped cells are activated in the same area. In theory, the cells would be activated at exactly one pixel. However, in any realistic scenario those cells will be grouped into a nearby region.

8.3. Curve Interpretation

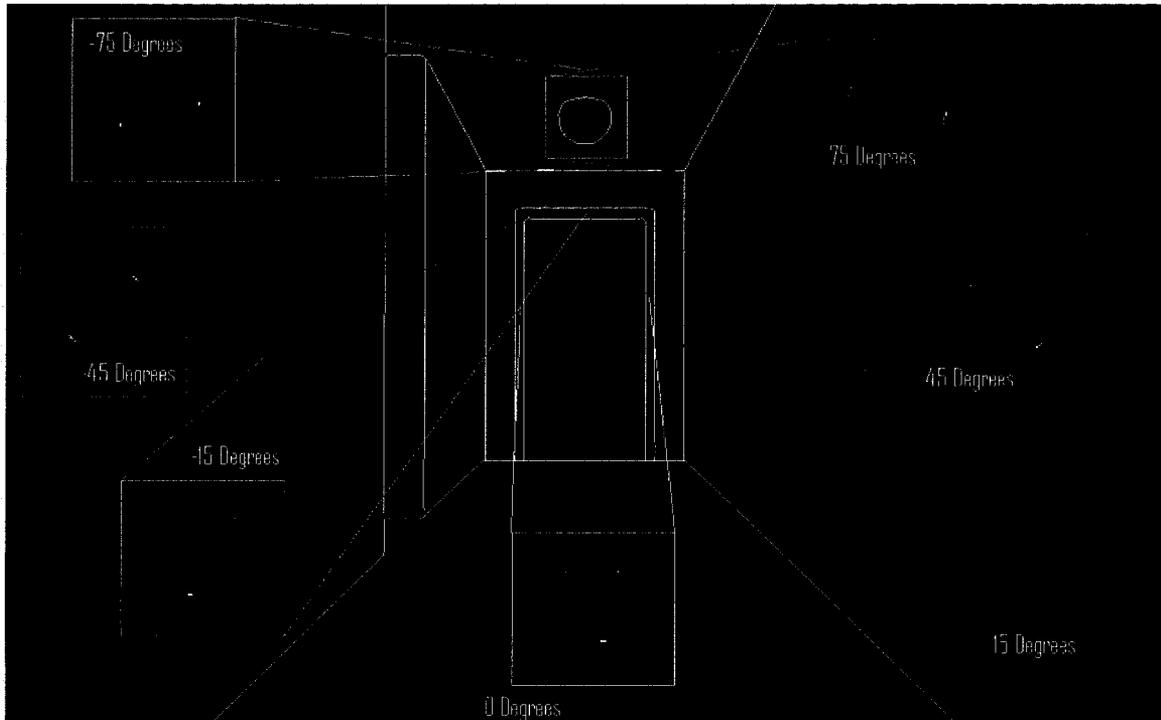


Figure 36 – Activation Around a Curve. Each Complex Cell identifies the regions of a curve that match its target orientation. Nearby edge segments will tend to be related to each other by neighbor orientation End Stops. Rather than seeing several joints around the edge, we tend to see the joining of neighbor cells as one continuous curving edge.

The cells react not only to edges and joints, but also to curves. Since a Complex Cell is essentially a first derivative operator, it reacts to segments of the curve that are close to the cell's orientation.

As was already pointed out, there is some overlap in the activity between adjacent cells. This has a consequence when dealing with curves; adjacent End-Stopped cells not activate at a single pixel. Instead, in many circumstances they will tend to activate along the curve but potentially some small distance away. They will activate on the curve because End-Stopped cells are trained to activate when the pixel being tested has a strong output value.

8.4. Off-Angle Orientation Interpretation

As discussed previously, adjacent Complex and End-Stopped cells are trained to have overlapping regions of activation. As a consequence some edges will be duplicated in the cell outputs. That is to say, the edge will show up in adjacent output maps.

Consider an edge with a 23 degree orientation (seen in Figure 37). This edge is in the Joint Activation Region between the 15 and 30 degree oriented cells (22.5 degrees being the exact middle). As expected, both cells find this edge. Careful examination of the 15 and 30 degree cells

show that the edge and its boundaries have been identified by all cells.

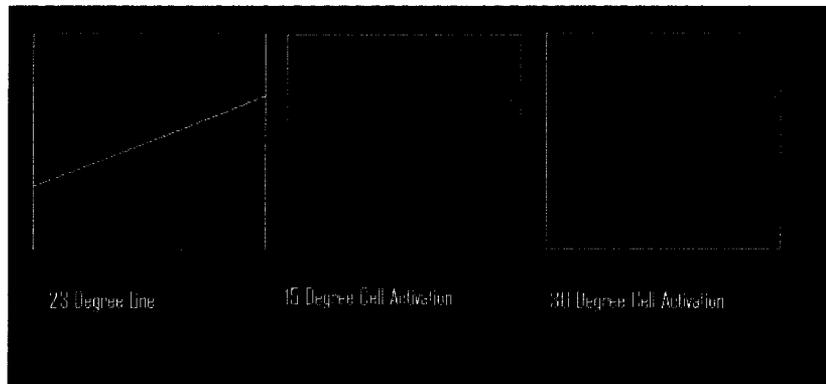


Figure 37 – 23 Degree Edge Triggers Multiple Cells. An edge that has a twenty three degree orientation will trigger reaction from both the fifteen and thirty degree oriented Complex Cells.

An edge with a 20 degree, while only separated by 3 degrees from the previous example, is processed quite differently. Its orientation falls in the 15 degree cell's Activation

Region and the 30 degree cell's Partial Activation Region.

In this case, the 15 degree cells properly identify the edge. In

contrast, the 30

degree cells only

find small portions of the edge. Dashed lines are produced as the Complex Cells react to

some portions of the line and not to

others. This in turn triggers the End

Stopped Cells to activate all along

the edge as well.

Taking a closer look at this partial

activation, it can be seen that the

cells actually find many small line

segments. These line segments line

up nicely; with each End Stopped

Cell matching an oppositely

oriented cell. The edges in a cell's

Partial Activation Region are represented by a chain of edges. The task of combining the

links of this chain into a single edge is left to future work.

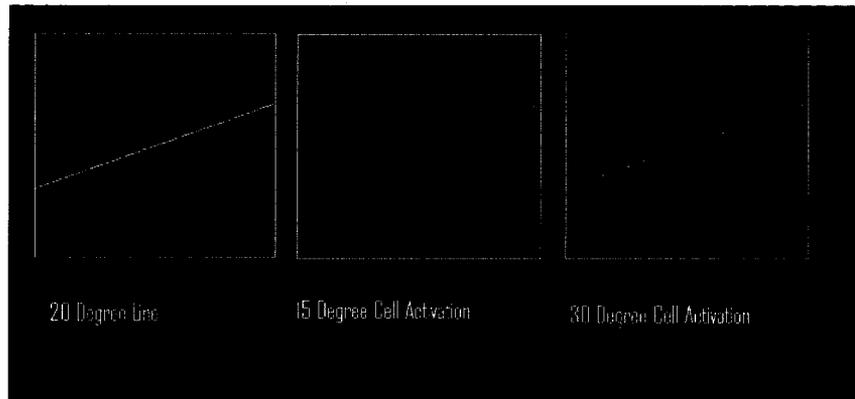


Figure 38 – 20 Degree Edge Partially Activates Cell. A twenty degree edge also triggers both the fifteen and thirty degree Complex Cells.

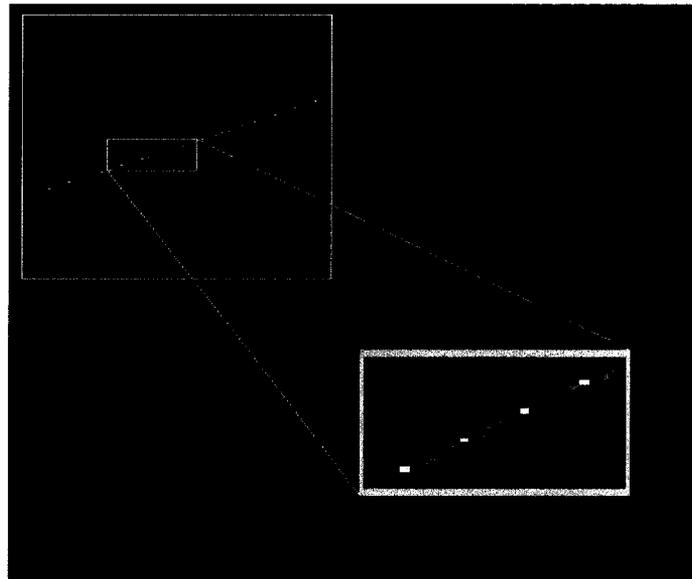


Figure 39 – Partially Activated Cells to Off-Angle Edge. The thirty degree cell only partially identifies the edge. It identifies small regions which must be stitched together by later processing via the matched End Stopped Cell activation.

The way that the cells react to edges in the Partial Activation Region has one interesting property. It displays the need for a mechanism to combine edge segments into a single larger edge. This behavior is very similar to behavior exhibited by humans. There is a well-known optical illusion where humans automatically fill in missing edge segments (see Figure 40) and combine edge segments into larger edges. It seems as though the human visual system is taking on similar issues those generated implicitly by the cells from this model.

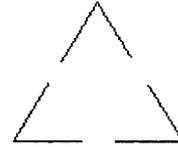
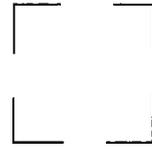


Figure 40 – Edge Reconstruction Optical Illusion. There seems to be a similar need to combine matched edges in the human system, as this occurs in this well known optical illusion.

8.5. Rounded Edges

Taking a first glance at the results generated by the hallway image, it is tempting to believe that there is a lot of noise in these images. However a closer examination of the cell outputs around these “noisy” spots reveals some interesting details. It can be seen in Figure 41 that the short line segments found in the -15 and -45 degree cell outputs are not actually noise, but reflect small changes in the orientation of the edges around the corner.

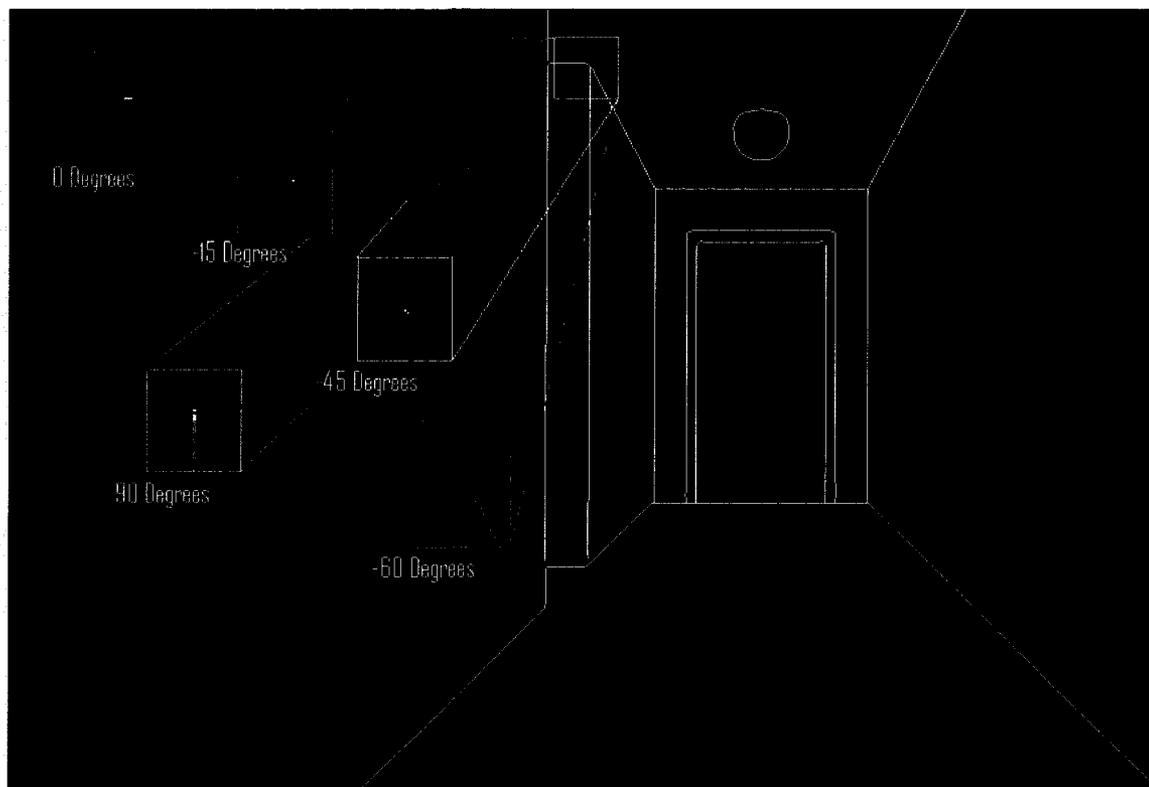


Figure 41 – Cell Reaction to Rounded Joint. Noisy output from the Edge Detection algorithm can cause the edge map to have a rounded appearance at the joints. To classify this, neighboring cells react around the curve. Non-neighbor cells then combine at their End Stops to form the joint.

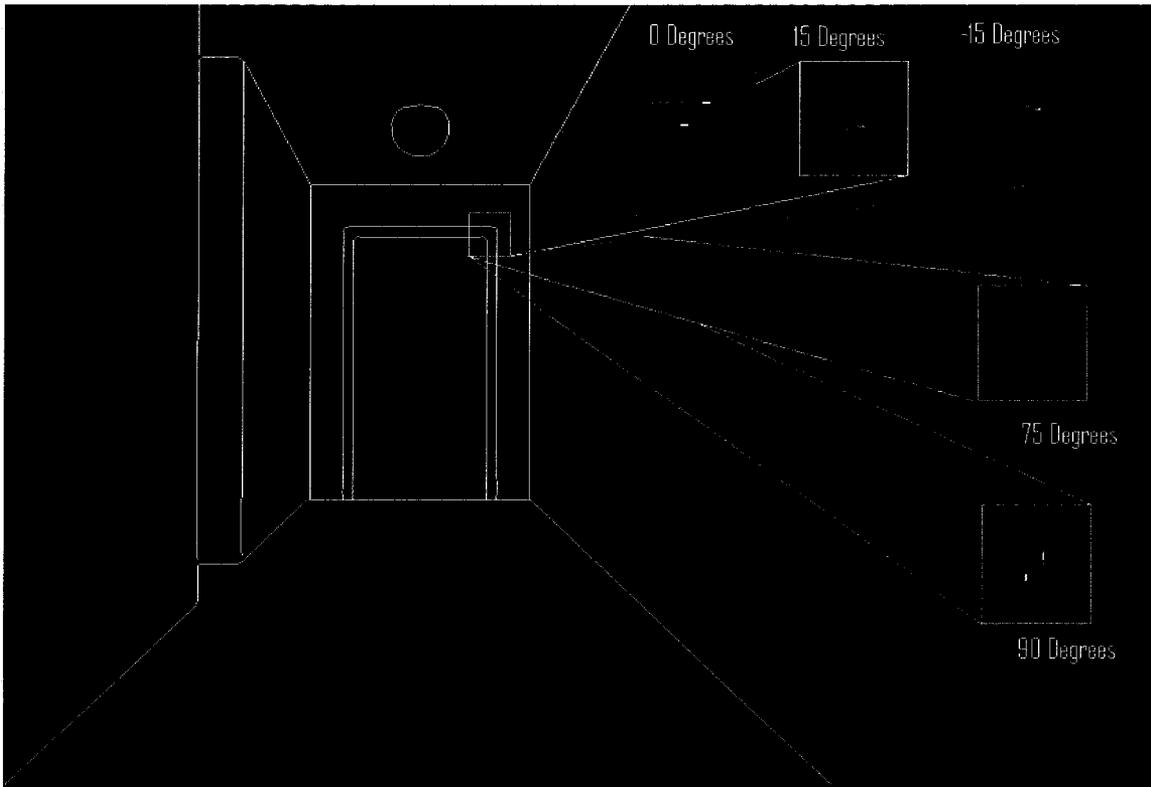


Figure 42 – Cell Reaction to Imperfect Edges and Joint. Another example where a rounded corner is interpreted as neighboring cell outputs form a single curved edge. These rounded edges join together with non-neighbor End Stopped Cells.

The region around the door frame displays similar “noise”. That is to say, small line segments displayed in the cell output actually reflect small imperfections in the underlying edge map. Careful study of the edge map will reveal that the lower portion of the frame has a small deformation causing the edge to bend slightly upwards. This deformation is matched by a small line segment detected by the 15 degree cells. Examination of all such noise in the cell output can be matched to small deformations in the underlying edge map (including the rounded edges seen in the previous example).

It seems as though the mechanism used to stitch together line segments into a longer single edge must be expanded to search neighbor cells rather than just output from a single cell. That way, the small rounded and jagged sections will be properly incorporated into the larger edge to which they correctly belong. Another way of stating

this is that an Edge Recognizer must track segments around nearby cells, and not just react to a long reaction to a single orientation.

8.6. *End Stopped Cells and Joint Identification*

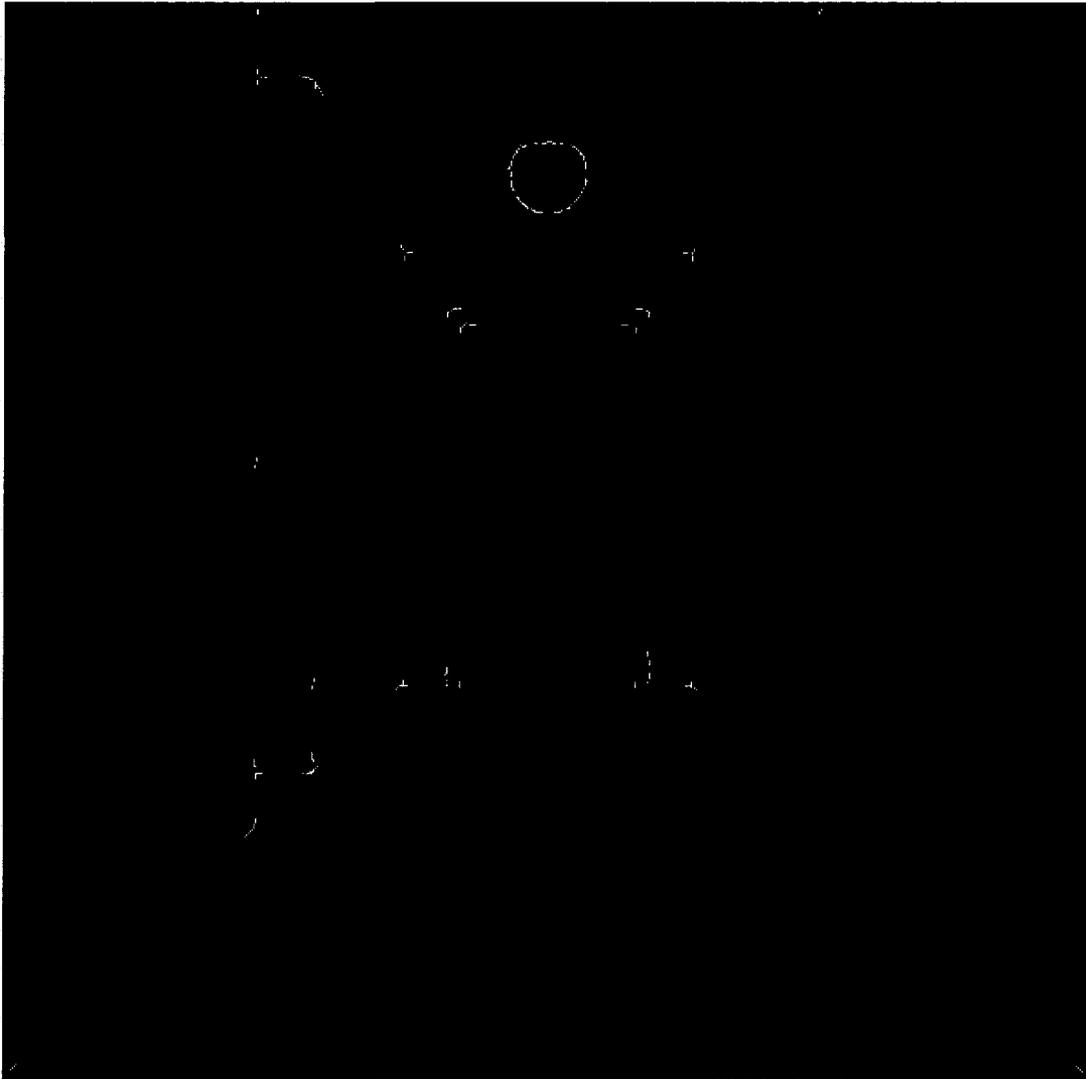


Figure 43 – Location of End Stopped Cell Activation. It can be seen that each joint is comprised of a cluster of End Stopped Cell activation. In addition, End Stopped Cells are activated around curved areas, allowing for arbitrary curve tracking.

One of the reasons for embarking on the work in this thesis was the belief that the End Stopped Cells could be used as Joint Recognizers. It has already been shown that the End Stopped Cells do activate at junctions. However, there still has been no explicit

check to see how strong the relationship between joints and End Stopped Cell activation actually is.

The location of the End Stopped activation in relation to the original edge map is displayed in Figure 43. The relationship between End Stopped Cell activation and the joints in the edge map are immediately obvious. Every single joint has been clearly marked by the End Stopped Cells.

There are also a number of places where the End Stopped Cells have activated which are not joints. The non-joint activations of the End Stopped Cells can be categorized into two types.

First, the cells fire all around the curve (i.e., the light fixture). Since the End Stopped Cells are reacting to changes in the Complex Cell activation, and the circle has continuously changing orientation, there is considerable activation of the End Stopped Cells. It seems that End Stopped Cells serve multiple purposes; joint identification or aids to curve tracking are prime examples. The End Stopped Cells around this curve fire along a neighbor Complex cell, and close to that Complex Cell's matching End Stopped Cell. In contrast, End Stopped Cells at joints lie at almost exactly the same point as an End Stopped Cell from non-Neighbor orientations.

Another type of place where End Stopped Cells can fire is at places where there is a discontinuity on the original edge map. When there is a small gap on an edge, or that edge is somewhat uneven, the resulting breaks in the Complex Cell output will cause the End Stopped Cell to activate. It has already been mentioned that there will have to be a

mechanism to chain together continuous edges that have been identified. This chaining of small edge segments will remove these sorts of the End Stops from consideration. It should be noted that the chaining together of small edge segments into one larger edge is very similar to the curve tracking already discussed. The only real distinction is that chaining is along the same orientation whereas curve tracking is along neighboring orientations. It seems likely that any mechanism to do one will be the same mechanism used to do the other.

The conclusion of this study is that End Stopped Cells can be used for two purposes. If it can be matched to another End Stopped Cell with a nearby (either a neighbor or exactly opposite) orientation, then the End Stopped Cell indicates that the line segments should be chained together to form one coherent object (either a straight line or a curve).

However, if the End Stopped Cell can be matched to one or more distant End Stops then the End Stop could represent a joint.

In the case of the rounded corners discussed previously, both of these features can be seen. Each edge coming into the joint can be chained to smaller edge segments, causing a rounded appearance. Now, these extended edges have non-neighbor End Stops (from the ends of the short segments added to the edges) which fall at common points which represent a joint.

9. Future Work

The research in this thesis has focused on the development of a logical hypercolumn; the building blocks from which other vision systems could be built. However, this is clearly only the first step in building a robust vision system.

The first work to be done should be the development of cross-hypercolumn cells. These cells would be used to join together and extend edges of the same or neighbor orientations which have been found in the immediate area. This would clean up a lot of the noise found around joints, as well as produce the optical illusion found when a single edge lies in a cell's Partial Activation Region.

The logical hypercolumns described here have been based on changes in grey scale image intensity. However, this is not the only feature which can be used to segment an image. Other types of features that could be considered would include colour, surface orientation, pattern and (if a video camera is employed) motion. There are other visual areas in the human visual system that process these types of features. A more robust study of these visual areas could produce better image segmentation than intensity alone. The orientation and end-stop results of all the different discriminators could then be combined to produce a single, reliable set of features which could be sent on to be processed by higher level processes (i.e., by the object recognition and localization streams of the visual system).

The next major piece of work that must be considered in the future is the development of the higher-level image interpretation streams. This includes both the Image Localization

and the Image Interpretation streams of the human visual system. This would require a detailed study of how these systems work for humans, so a detailed model of these parts of the system could be built and implemented.

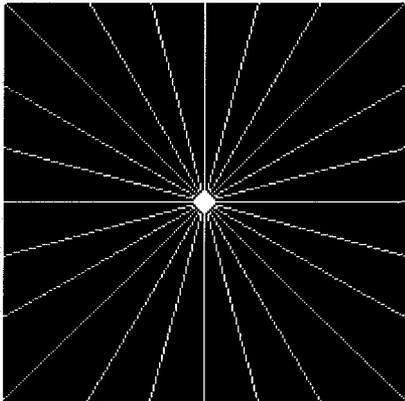
Finally, the human system has a large amount of feedback built into it. The nature of this feedback could provide valuable information to earlier portions of the vision system; particularly if the system is processing motion (as the change from image to image over time could be processed). Careful study of the mechanisms through which the current state of the vision system informs the processing of new information could provide some very interesting new mechanisms for future upgrades to the cells developed here.

10. Appendix

Following are the raw outputs from the Complex and End Stopped Cells when presented with a number of different testing examples. As has been done previously, the cells with matching orientations have their outputs displayed together for easier reference.

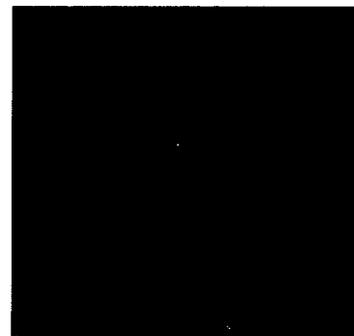
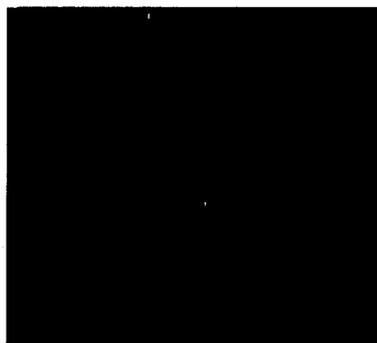
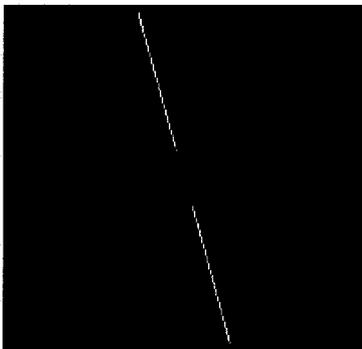
10.1. *Exact Target Angles*

This section displays the results of the Complex and End-Stopped Cells against an artificial edge map that represents lines that match exactly the target edges that the cells have been trained to react to.



10.1.1. **-75 Degree Results**

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



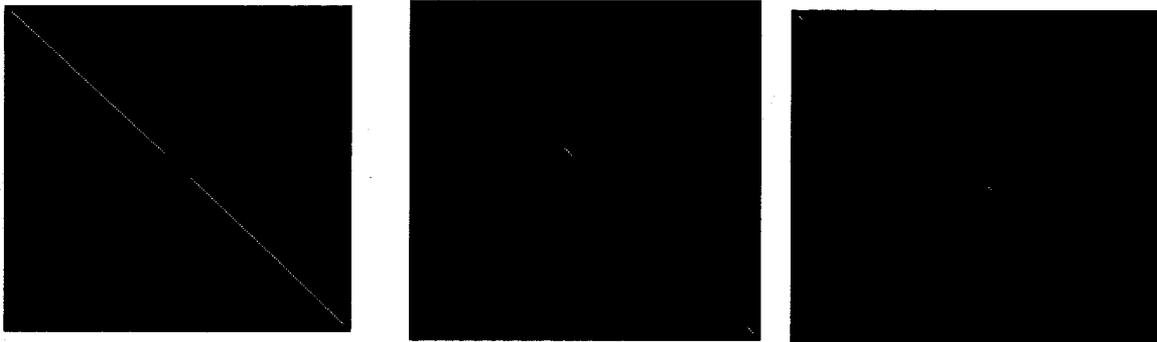
10.1.2. -60 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



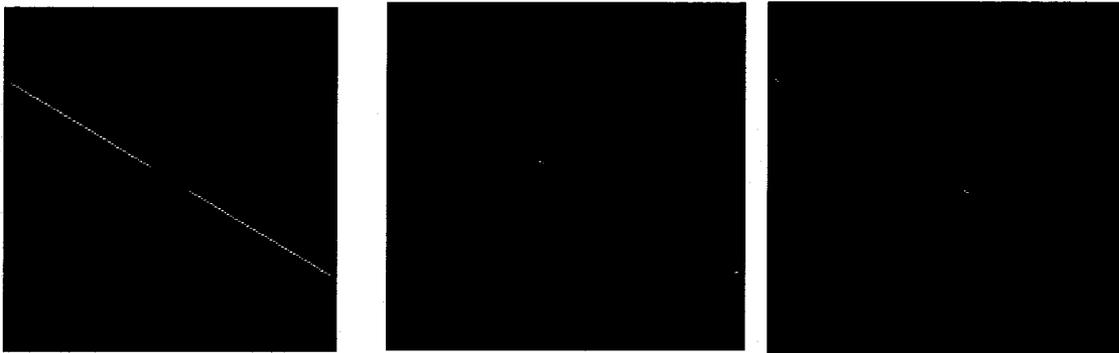
10.1.3. -45 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



10.1.4. -30 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



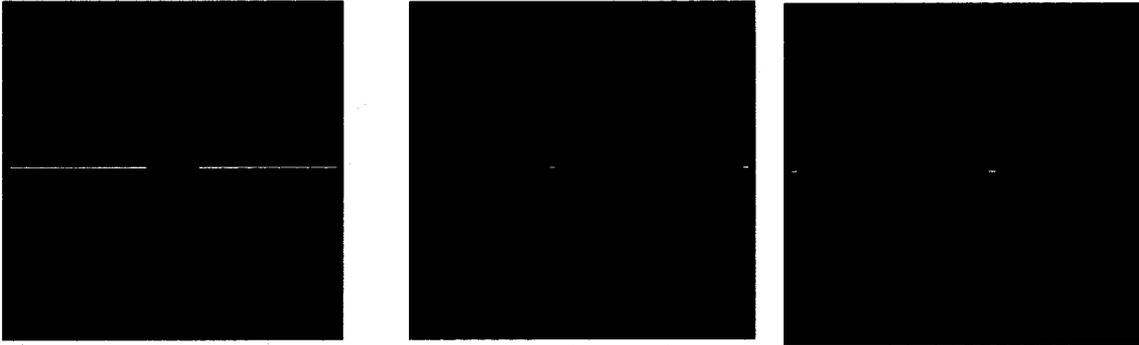
10.1.5. -15 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



10.1.6. 0 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



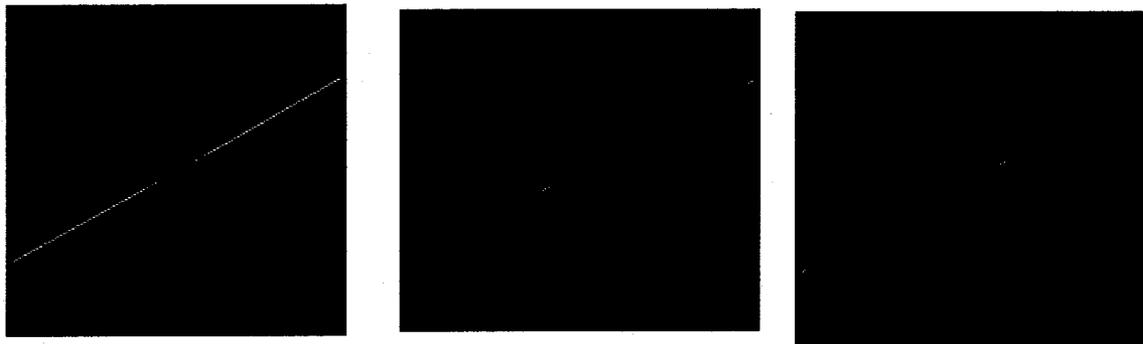
10.1.7. 15 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



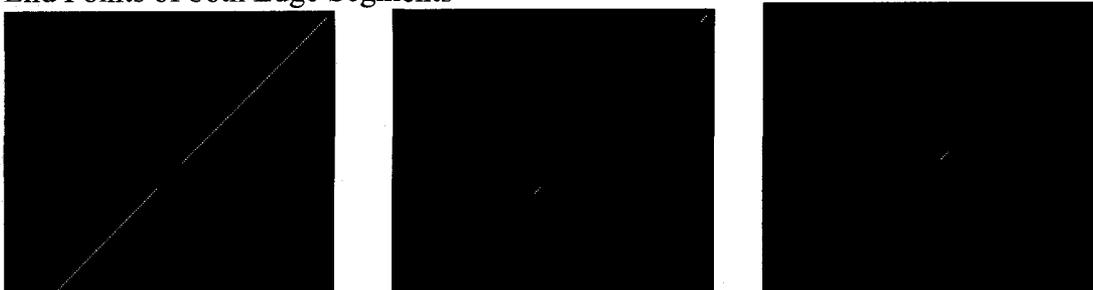
10.1.8. 30 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



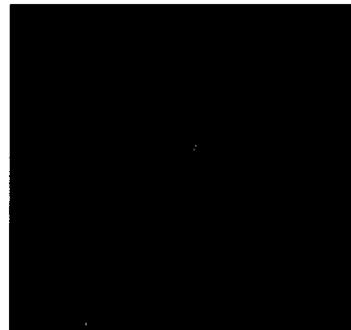
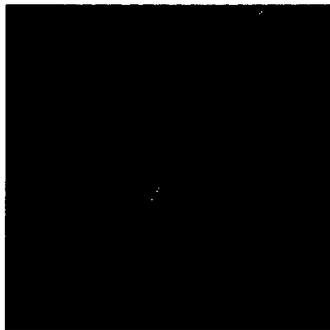
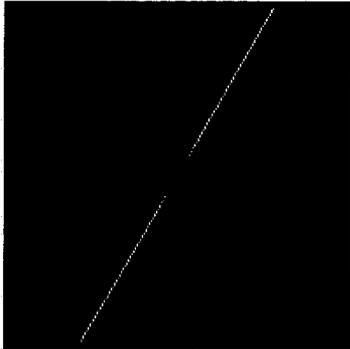
10.1.9. 45 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



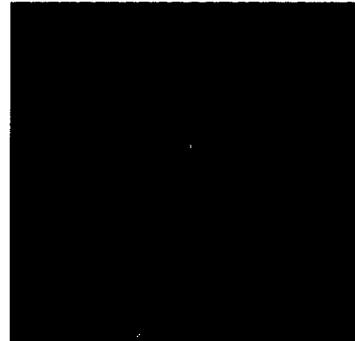
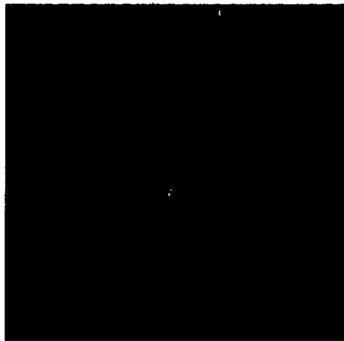
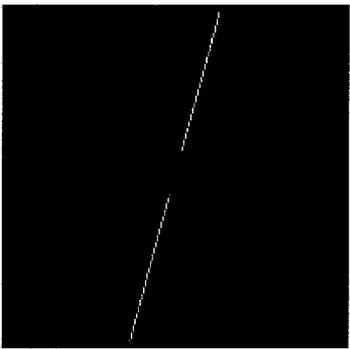
10.1.10. 60 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments



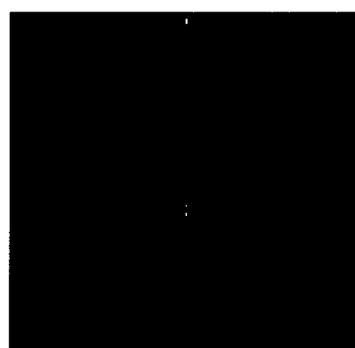
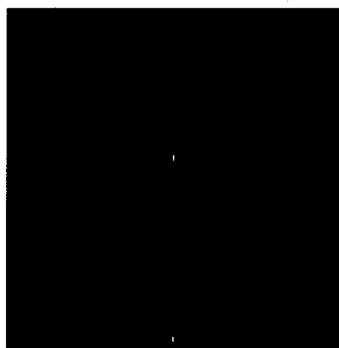
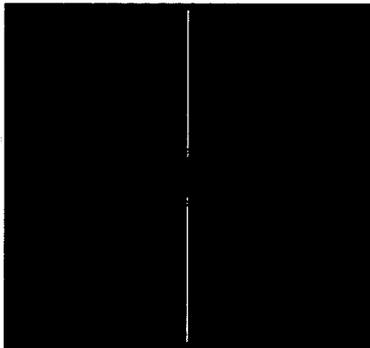
10.1.11. 75 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments

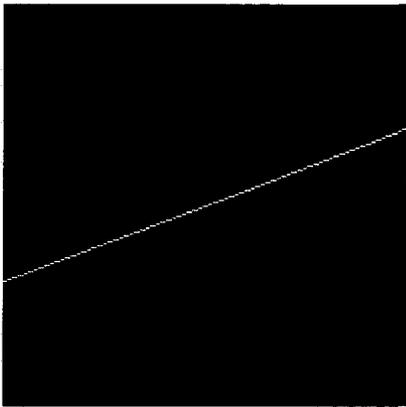


10.1.12. 90 Degree Results

Edge Correctly Found by Complex Cell, End Stopped Cells Correctly Identify Start and End Points of both Edge Segments

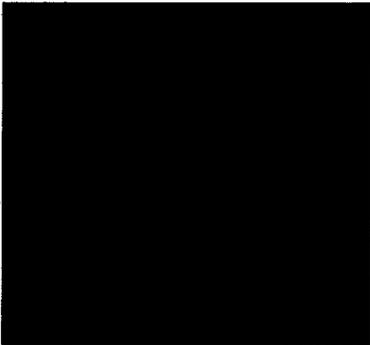


10.2. 20 Degree Line



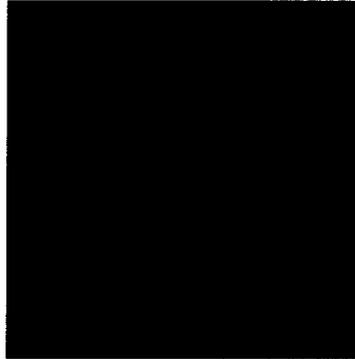
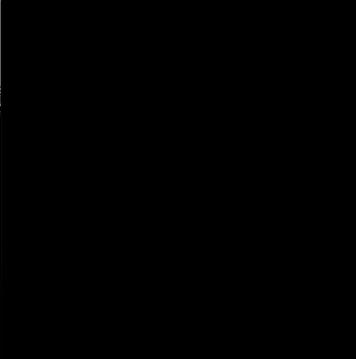
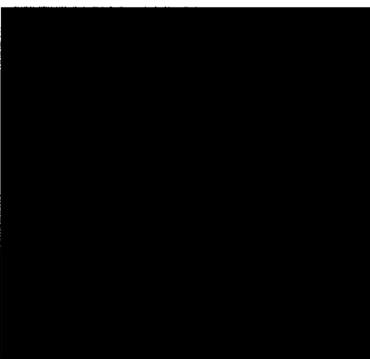
10.2.1. -75 Degree Results

No reactions to any cells, since there are no edges that match this orientation



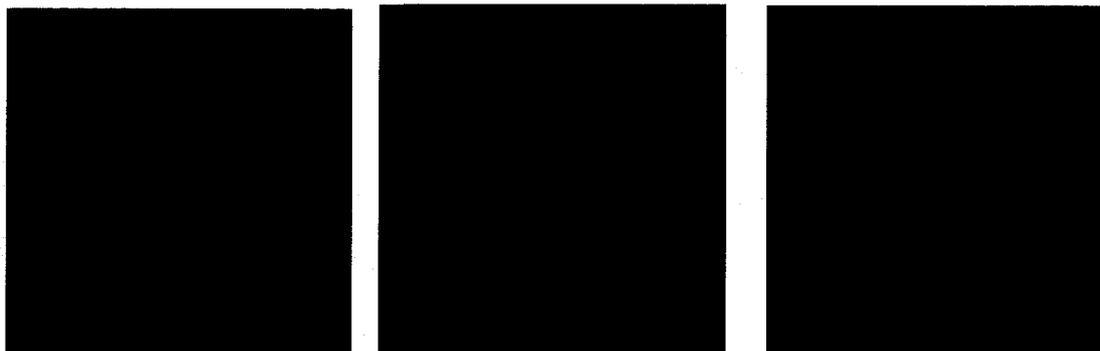
10.2.2. -60 Degree Results

No reactions to any cells, since there are no edges that match this orientation



10.2.3. -45 Degree Results

No reactions to any cells, since there are no edges that match this orientation



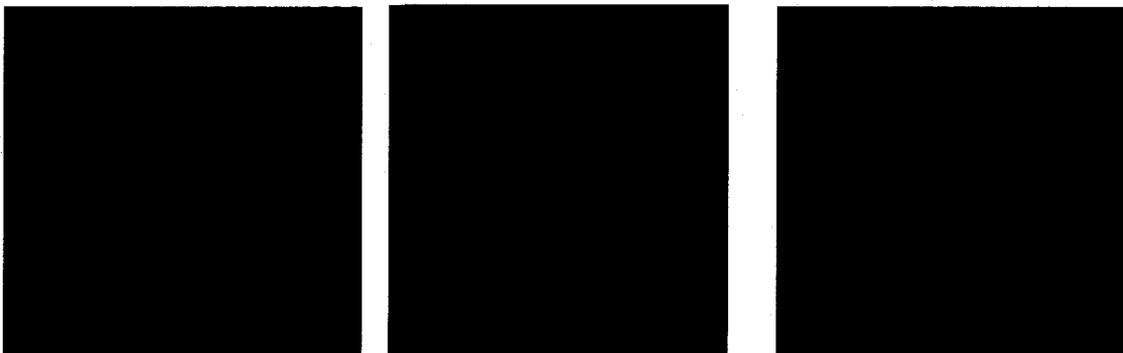
10.2.4. -30 Degree Results

No reactions to any cells, since there are no edges that match this orientation



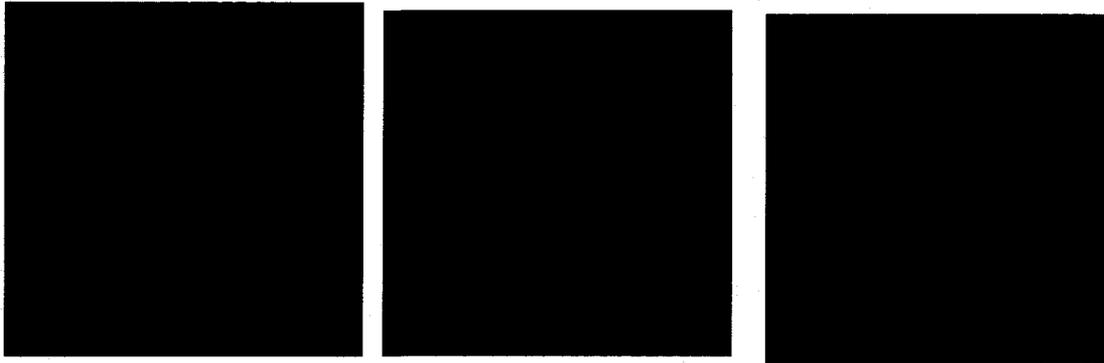
10.2.5. -15 Degree Results

No reactions to any cells, since there are no edges that match this orientation



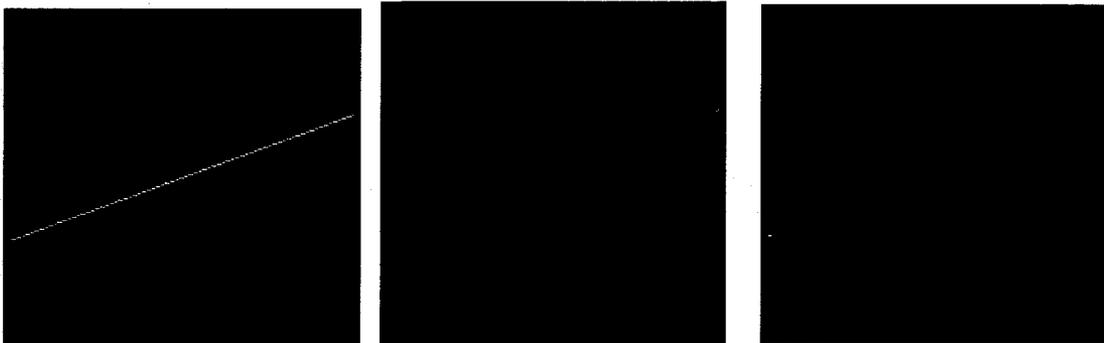
10.2.6. 0 Degree Results

No reactions to any cells, since there are no edges that match this orientation



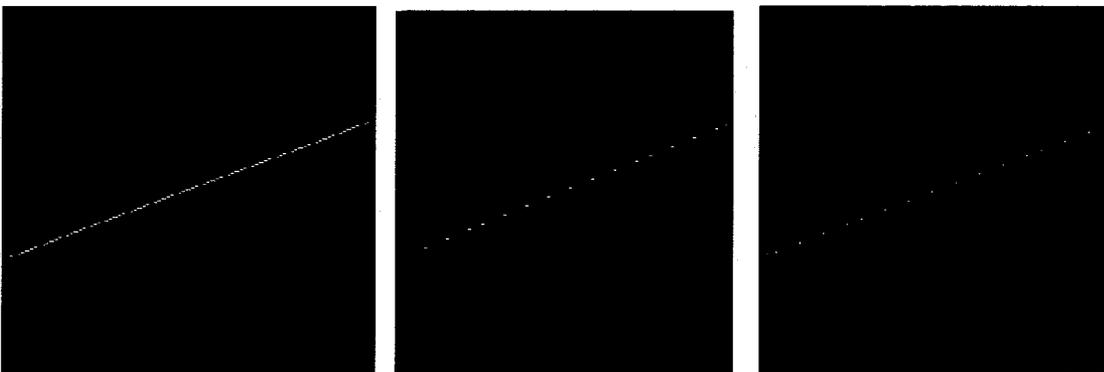
10.2.7. 15 Degree Results

These cells correctly identify the nearby oriented edge presented



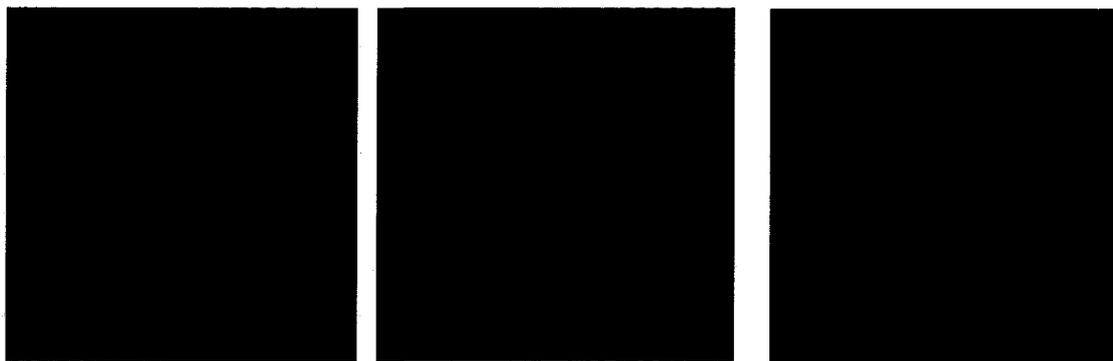
10.2.8. 30 Degree Results

These cells partially identify the nearby edge presented. Extra outputs to the End Stopped Cells are generated to stitch together the dashed output



10.2.9. 45 Degrees Results

No reactions to any cells, since there are no edges that match this orientation



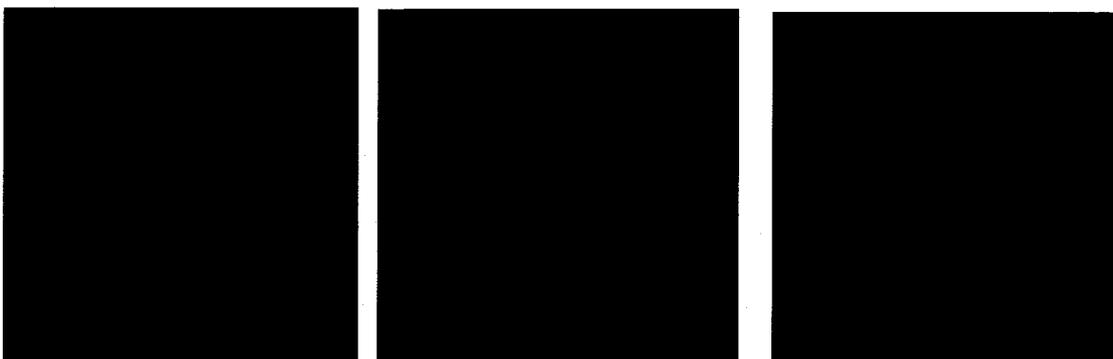
10.2.10. 60 Degree Results

No reactions to any cells, since there are no edges that match this orientation



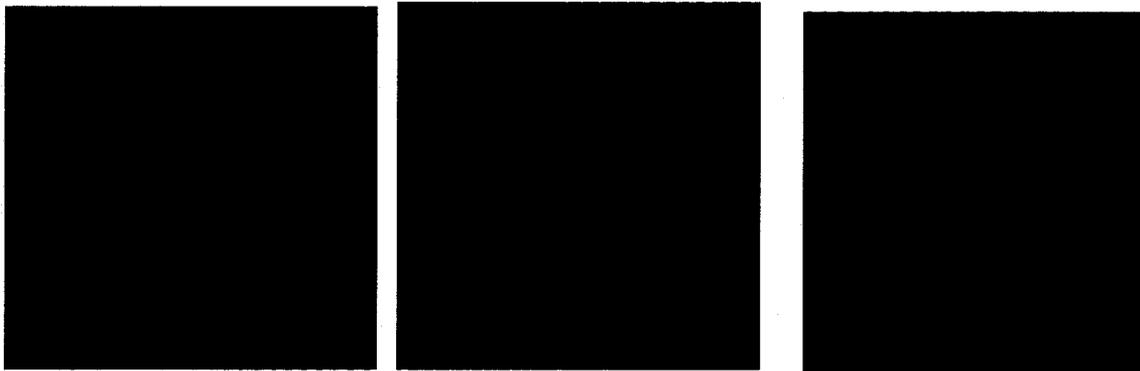
10.2.11. 75 Degree Results

No reactions to any cells, since there are no edges that match this orientation

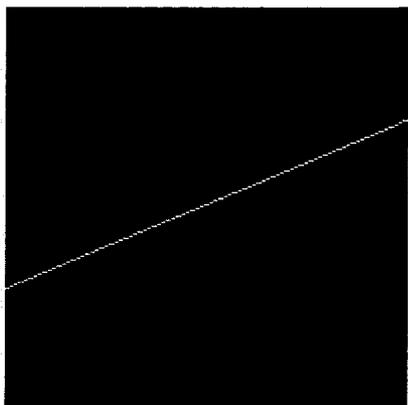


10.2.12. 90 Degree Results

No reactions to any cells, since there are no edges that match this orientation

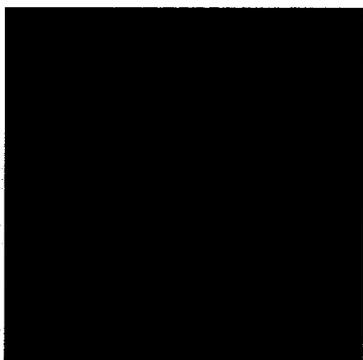


10.3. 23 Degree Line



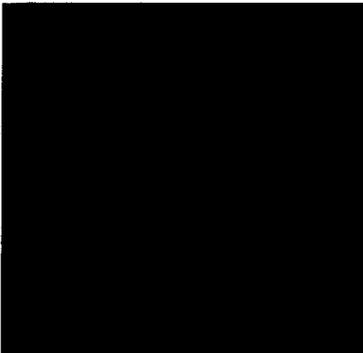
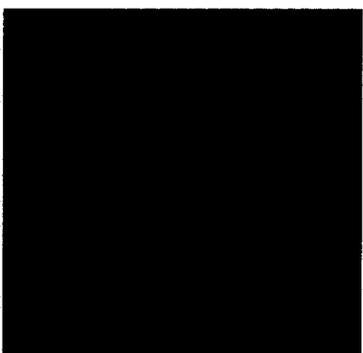
10.3.1. -75 Degree Results

No reactions to any cells, since there are no edges that match this orientation



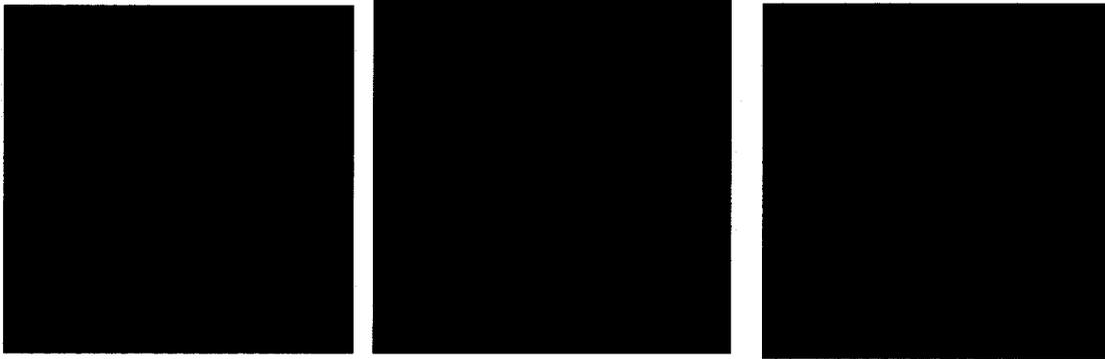
10.3.2. -60 Degree Results

No reactions to any cells, since there are no edges that match this orientation



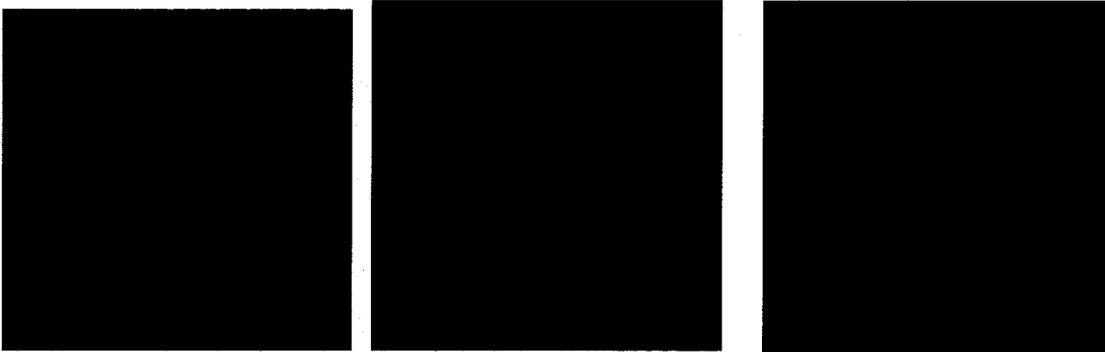
10.3.3. -45 Degree Results

No reactions to any cells, since there are no edges that match this orientation



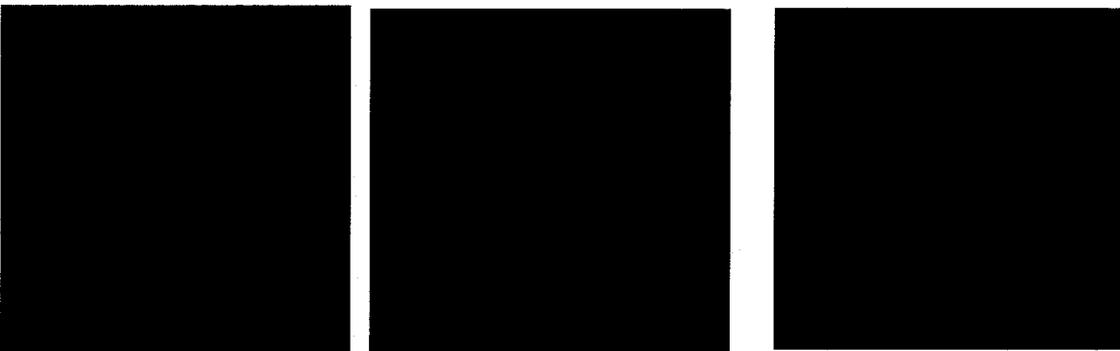
10.3.4. -30 Degree Results

No reactions to any cells, since there are no edges that match this orientation



10.3.5. -15 Degree Results

No reactions to any cells, since there are no edges that match this orientation



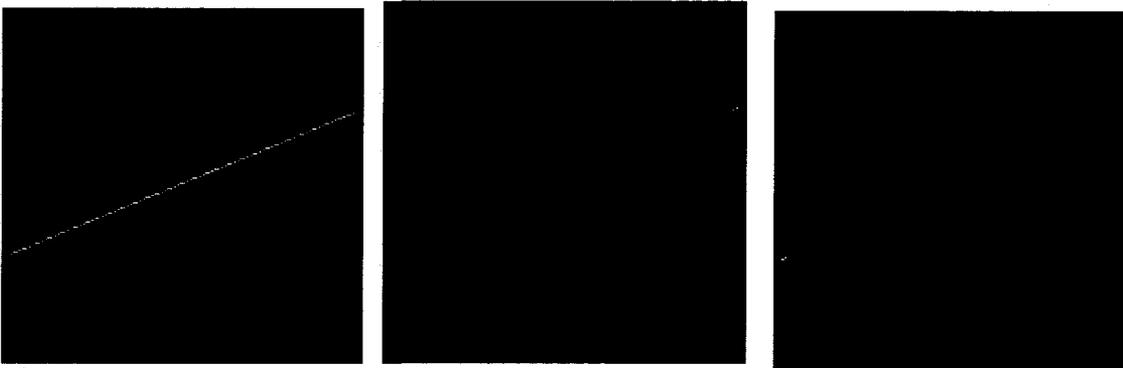
10.3.6. 0 Degree Results

No reactions to any cells, since there are no edges that match this orientation



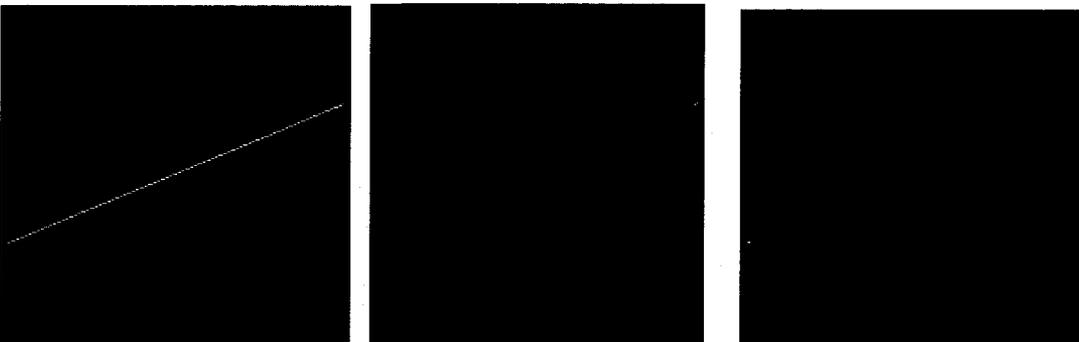
10.3.7. 15 Degree Results

These networks correctly identify the nearby angle



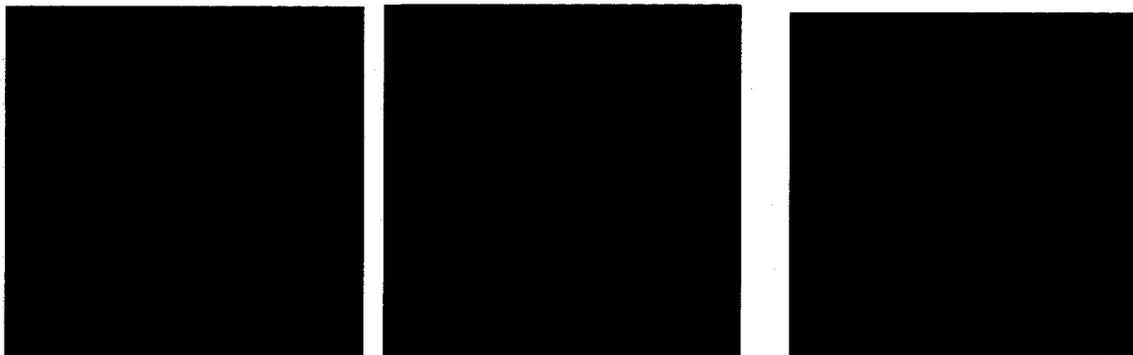
10.3.8. 30 Degree Results

These networks correctly identify the nearby angle



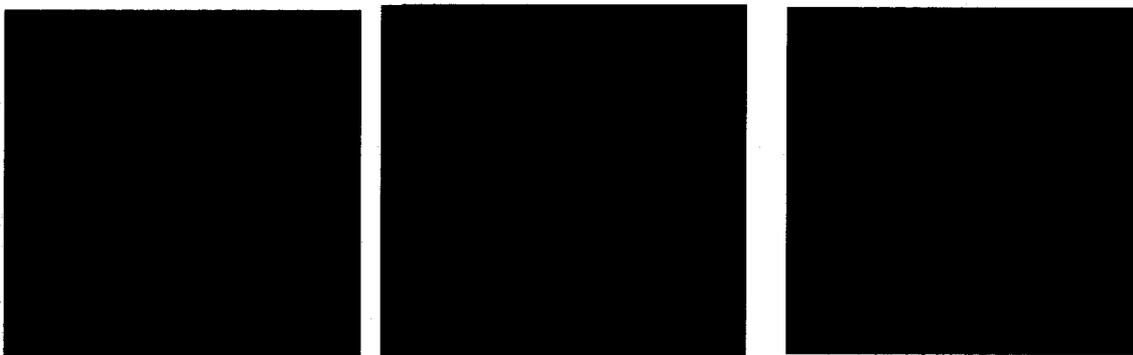
10.3.9. 45 Degree Results

No reactions to any cells, since there are no edges that match this orientation



10.3.10. 60 Degree Results

No reactions to any cells, since there are no edges that match this orientation



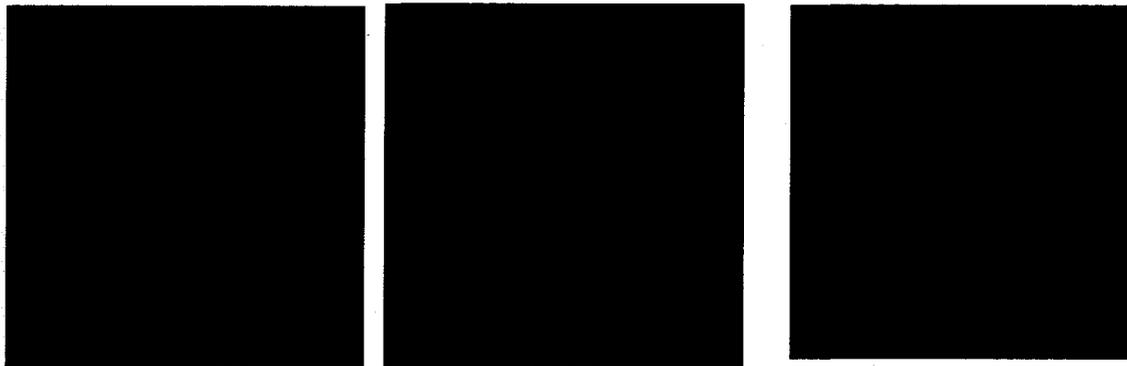
10.3.11. 75 Degree Results

No reactions to any cells, since there are no edges that match this orientation



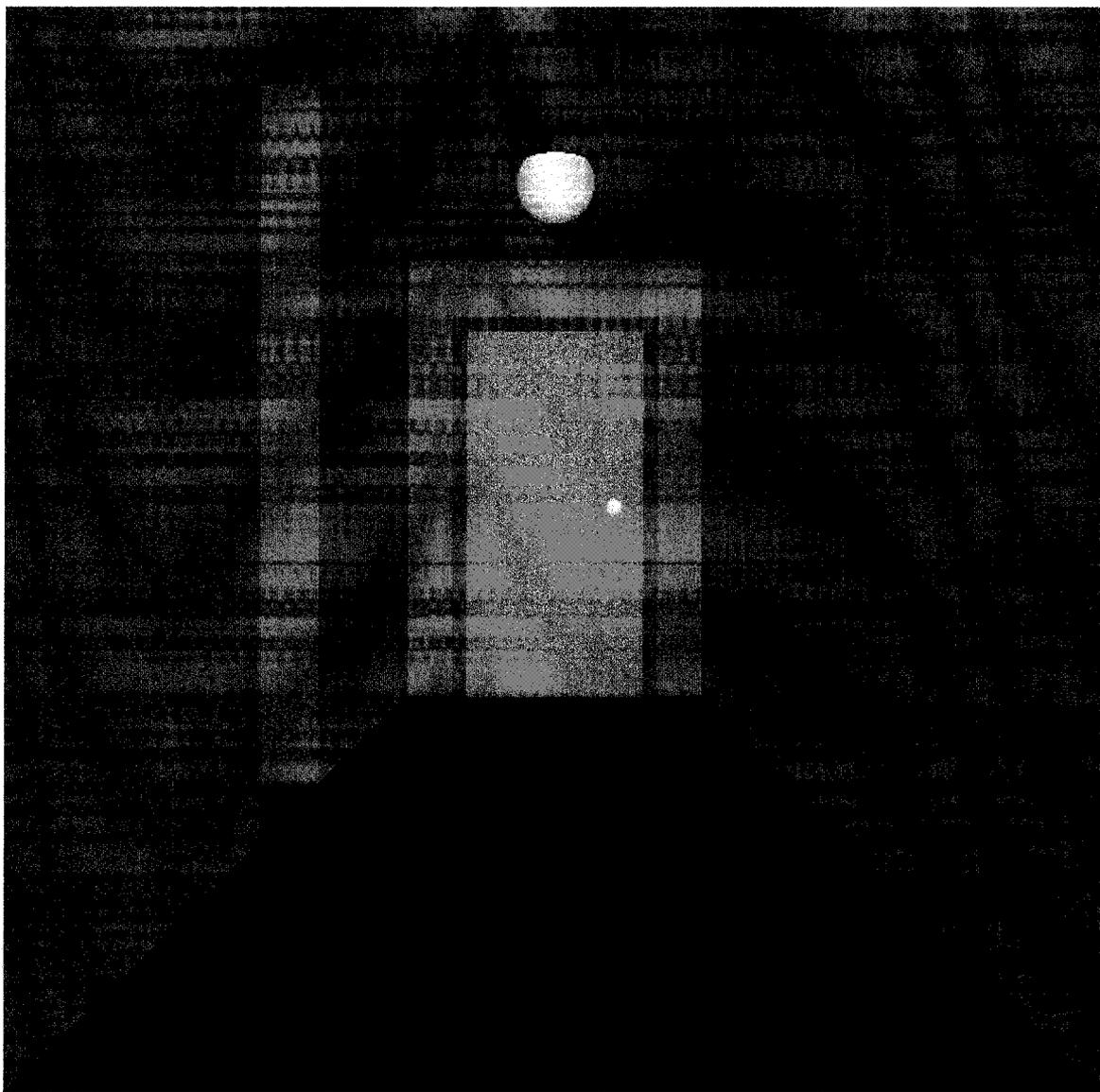
10.3.12. 90 Degree Results

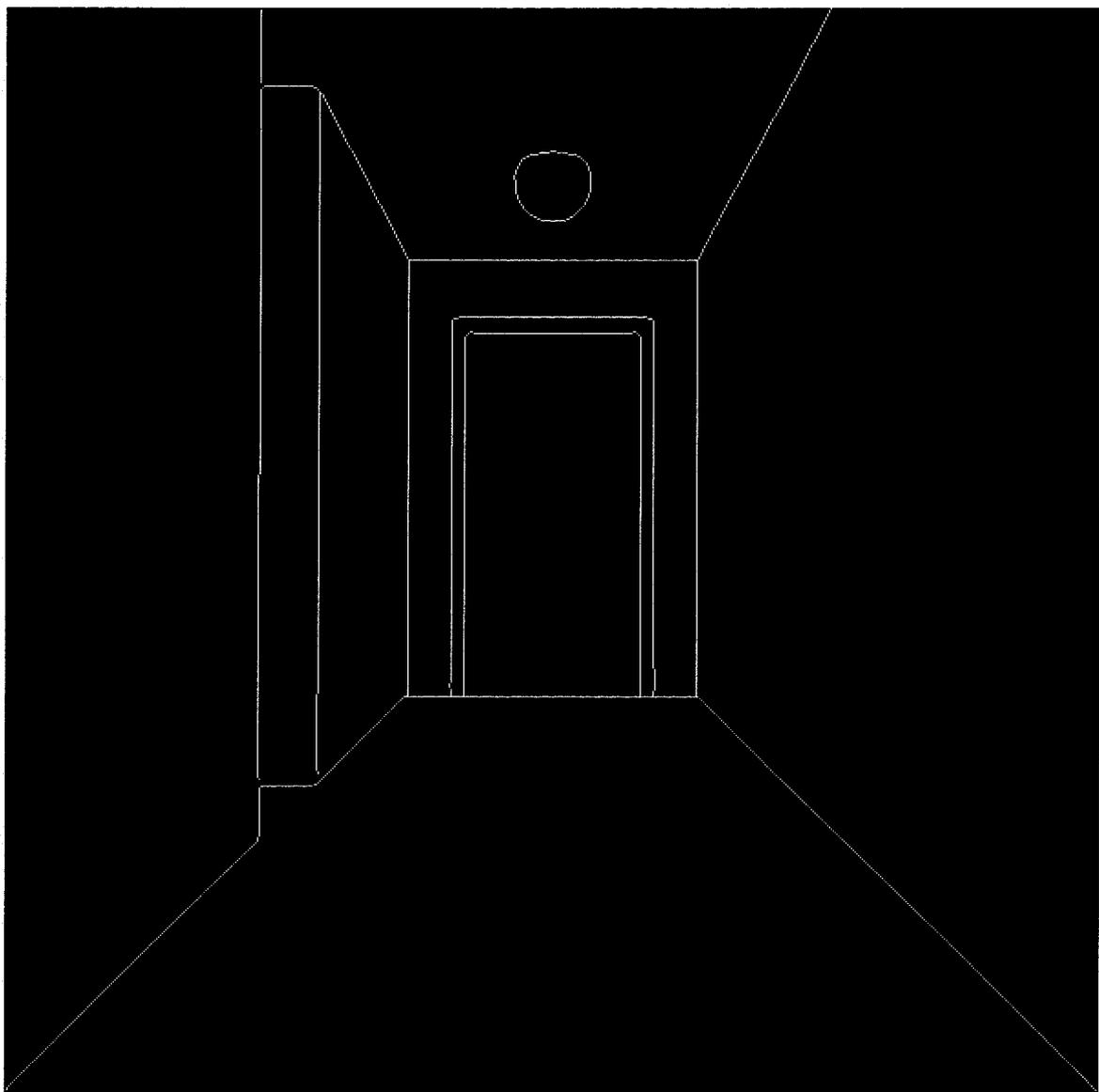
No reactions to any cells, since there are no edges that match this orientation



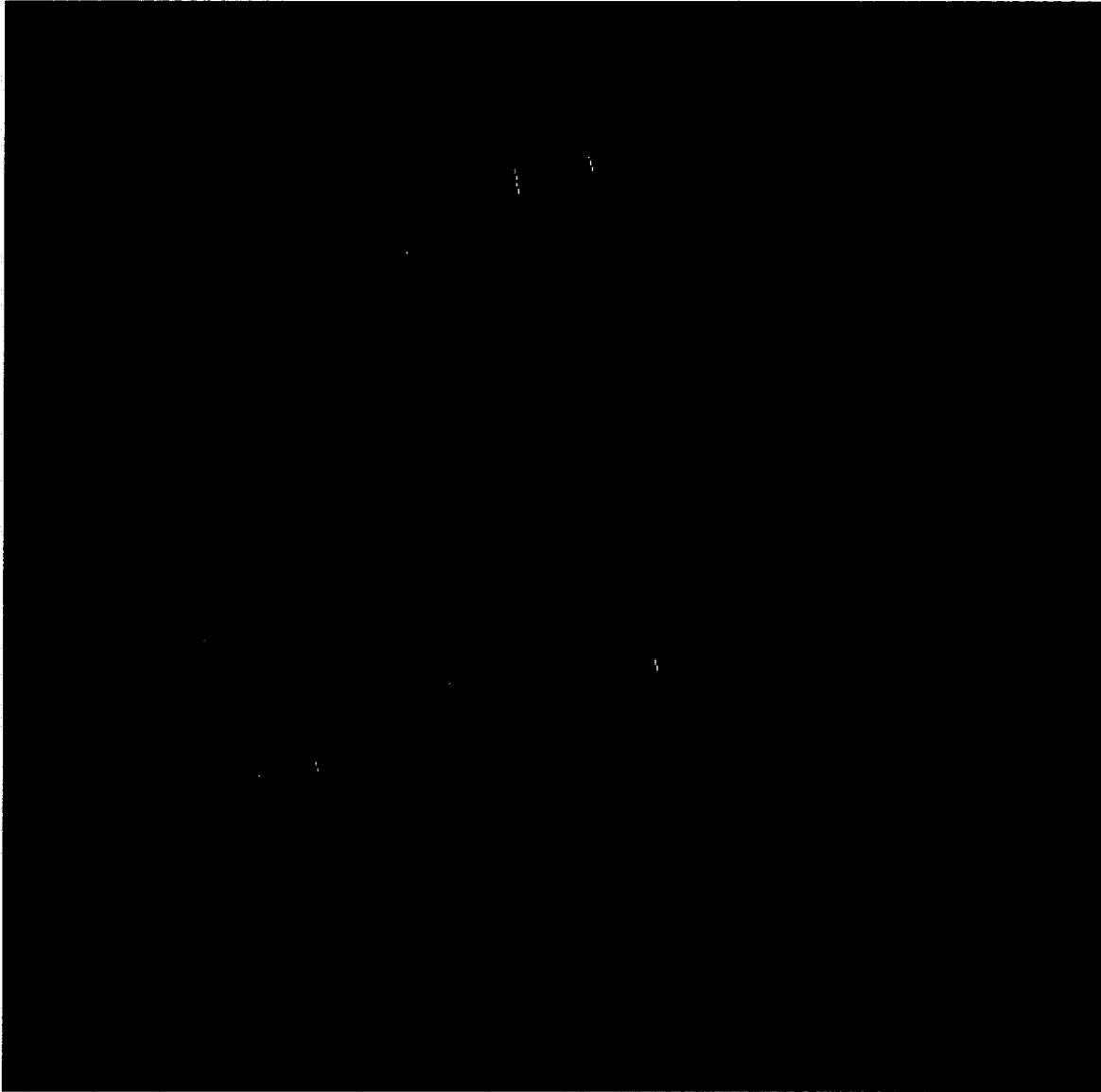
10.4. Hallway Results

The following image represents a more realistic test of the V1 Model. An image that is a closer representation of an image that a system could interpret is presented to the system. Preprocessing steps are applied to this image, and an edge map is generated, which is then applied to the Complex and End-Stopped Cells.





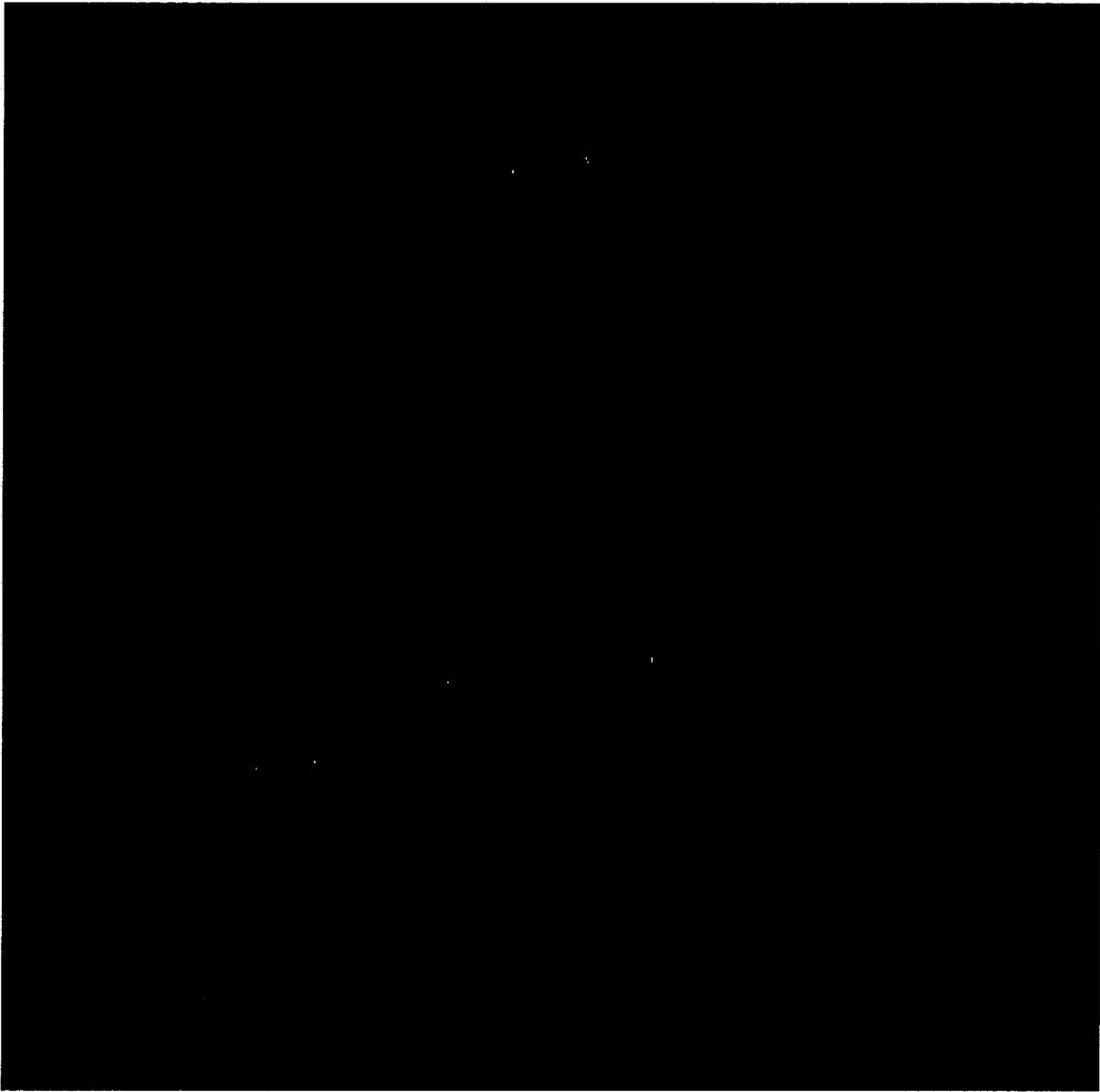
10.4.1. -75 Degree Results



Small line segments aligned in the matching orientation are correctly identified.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments.

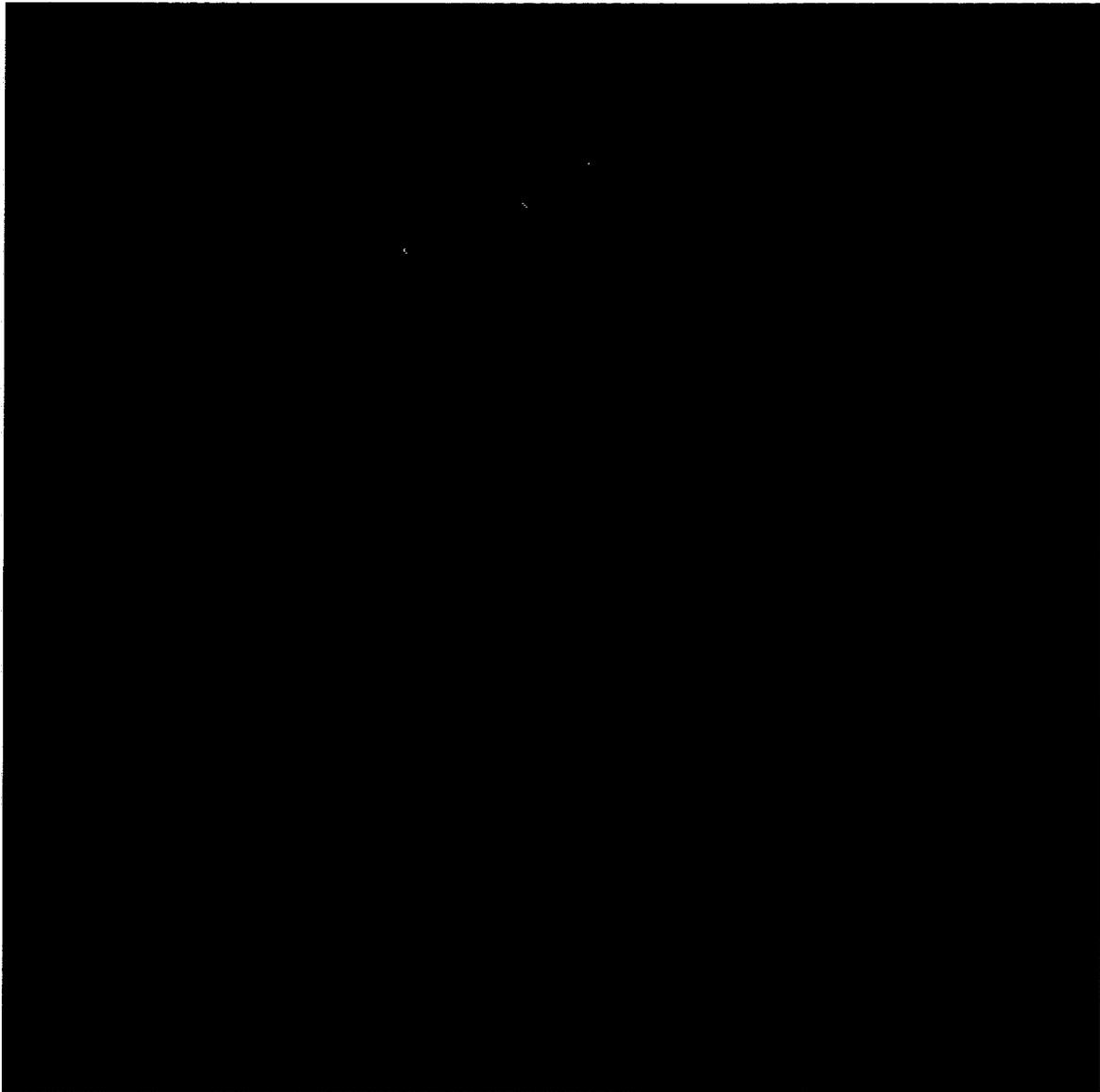


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments.

10.4.2. -60 Degree Results



Small line segments and one longer edge which is aligned in the matching orientation are correctly identified.

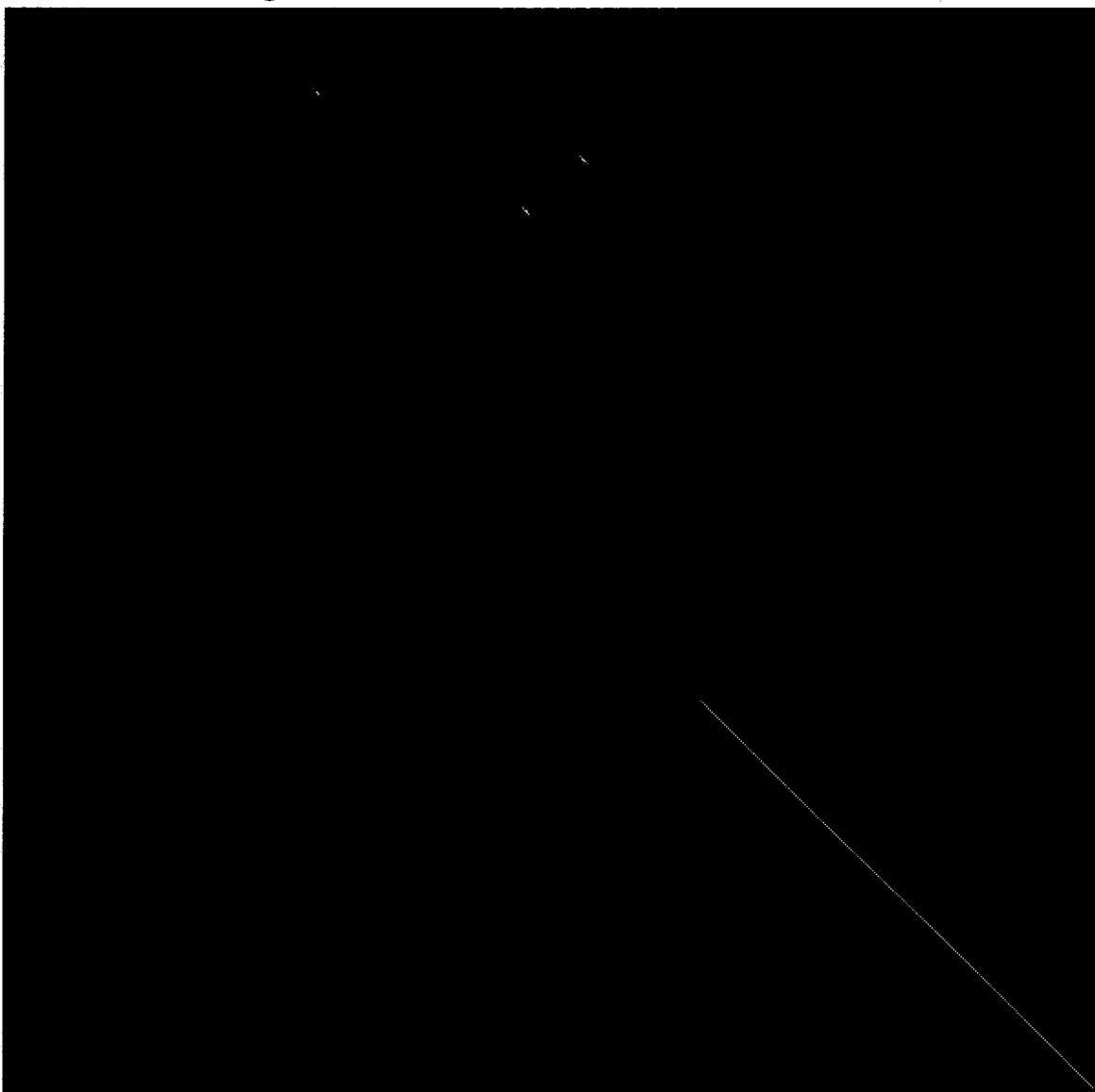


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edge.

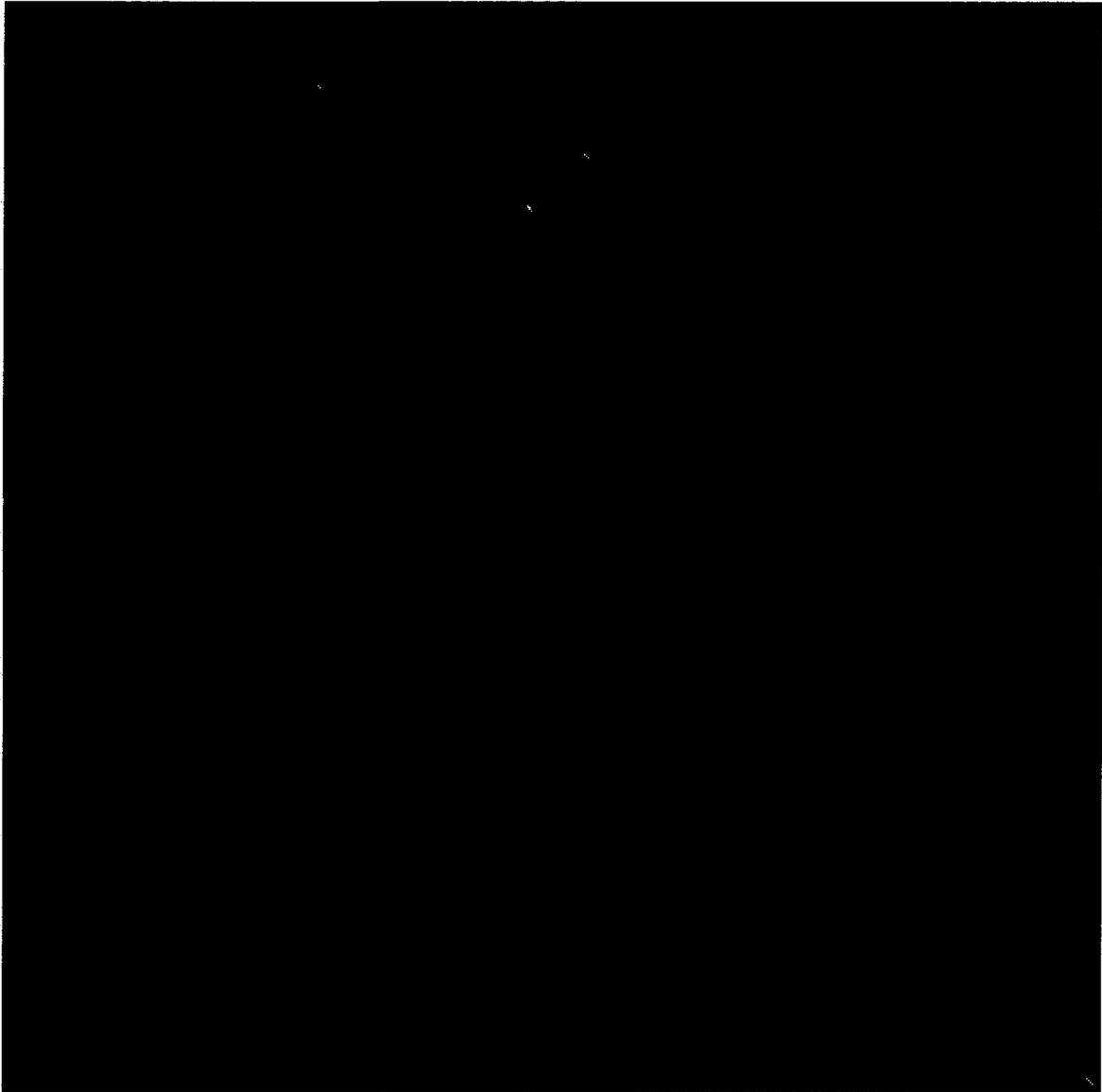


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edge.

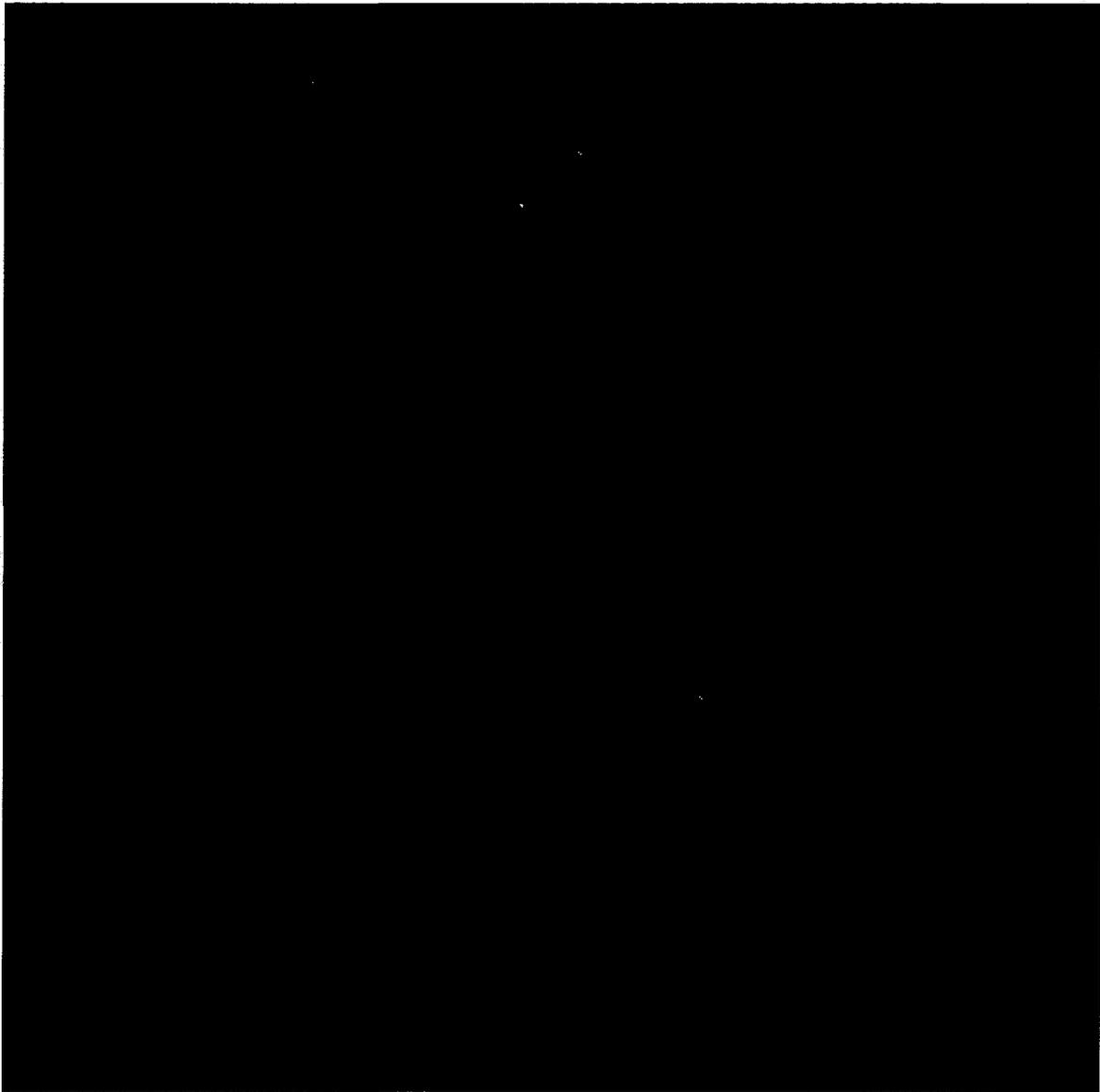
10.4.3. -45 Degree Results



Small line segments and a long edge which is aligned in the matching orientation are correctly identified.

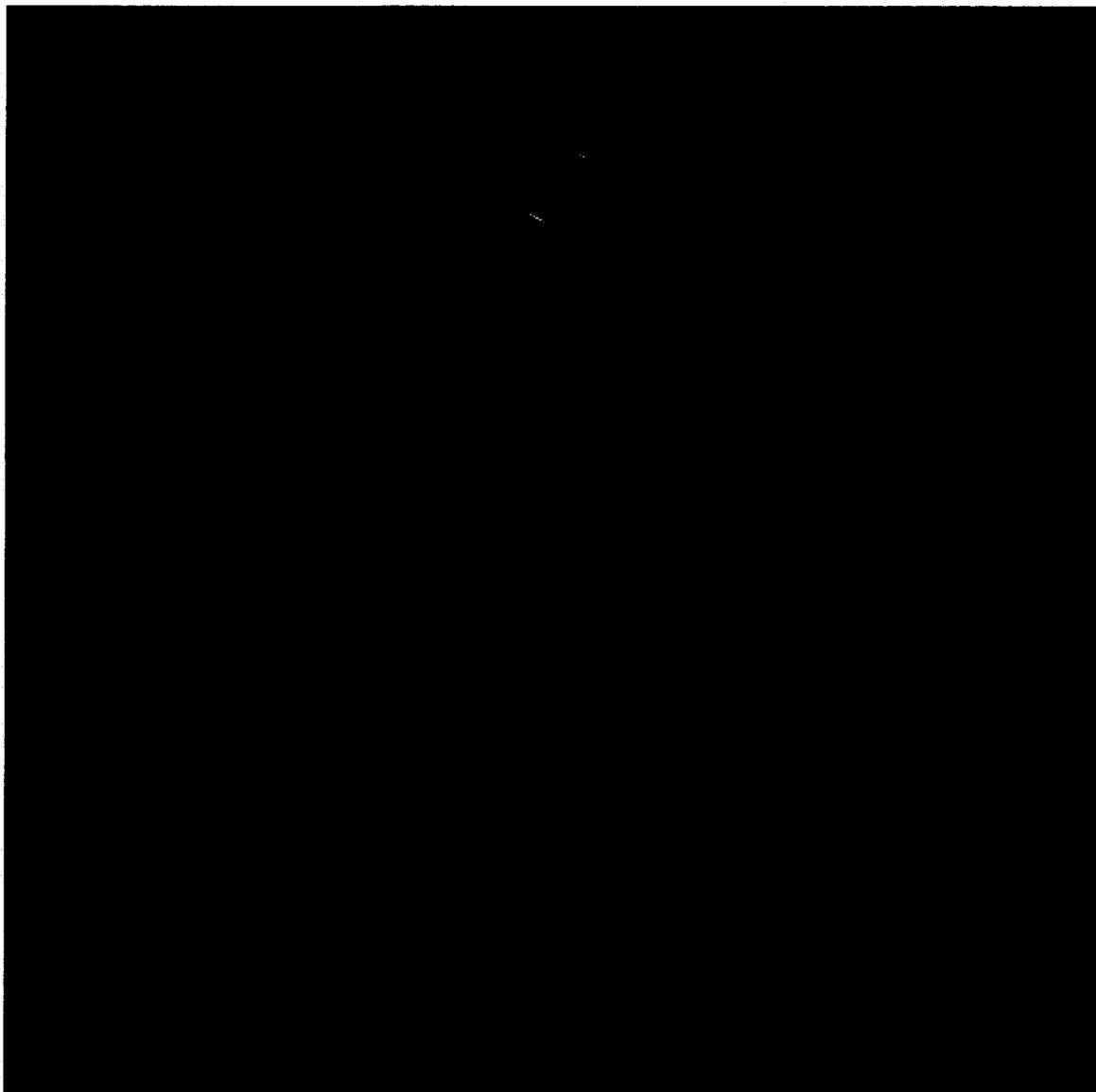


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edge.

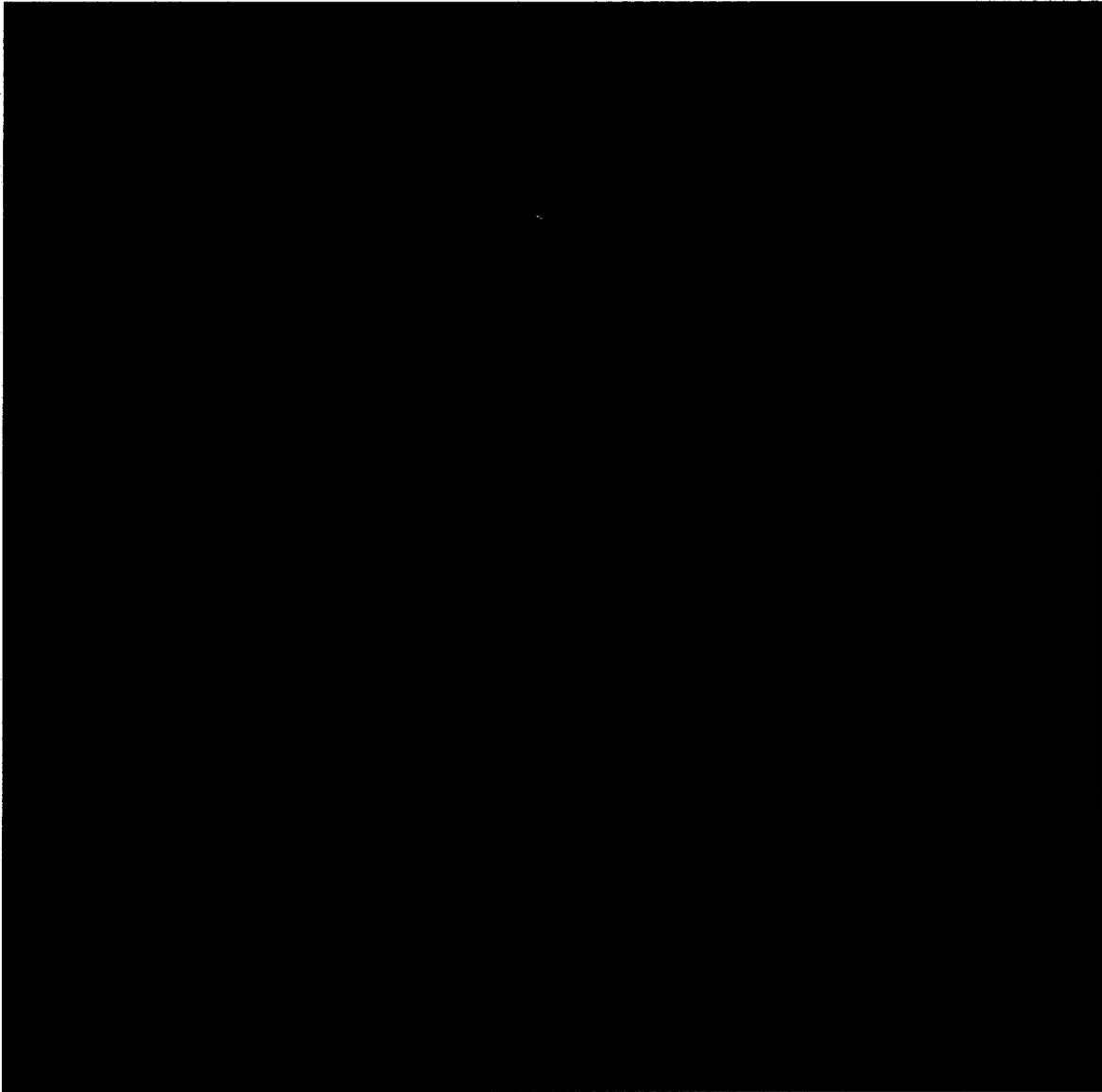


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edge.

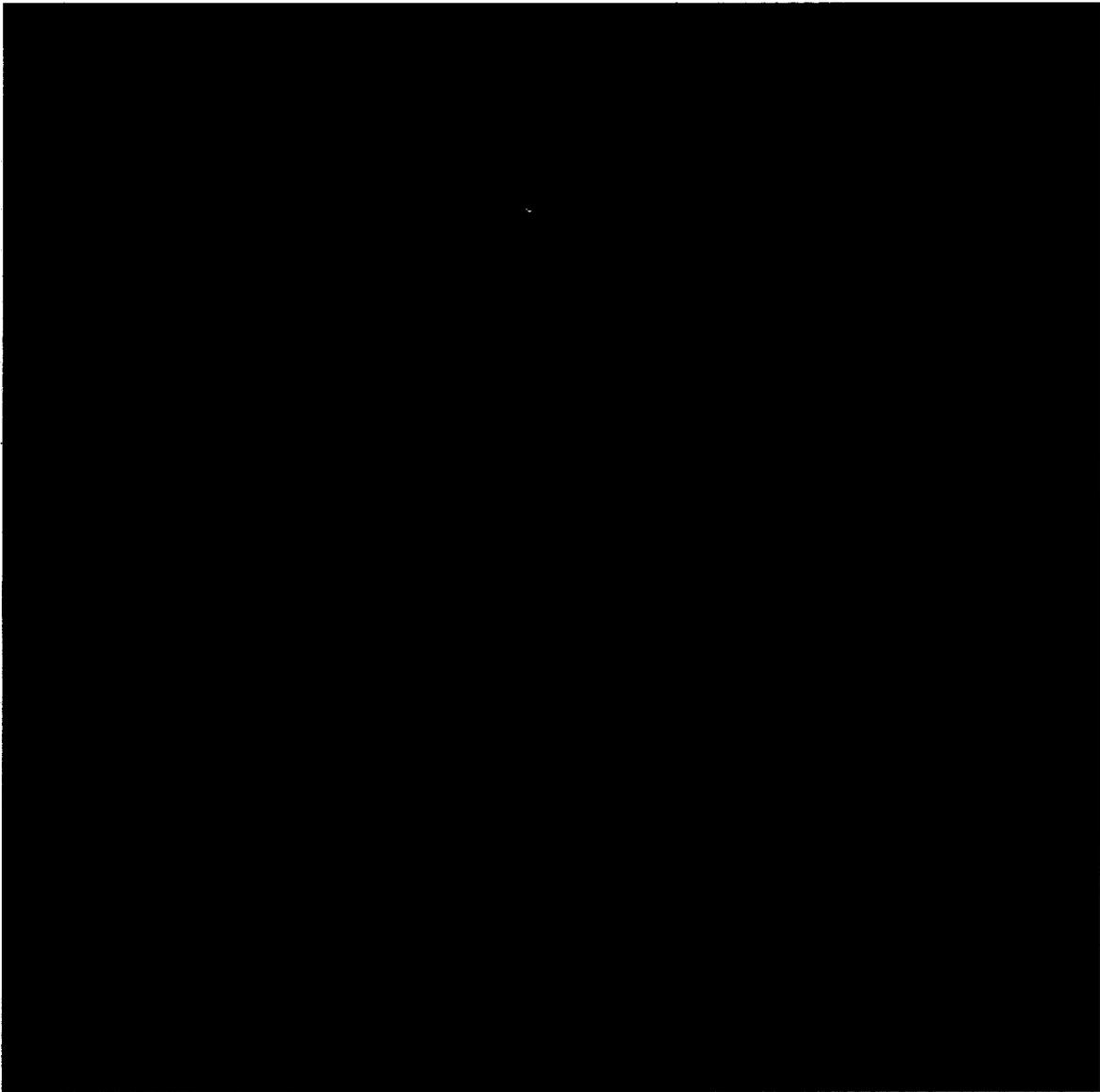
10.4.4. -30 Degree Results



Small line segments aligned in the matching orientation are correctly identified.

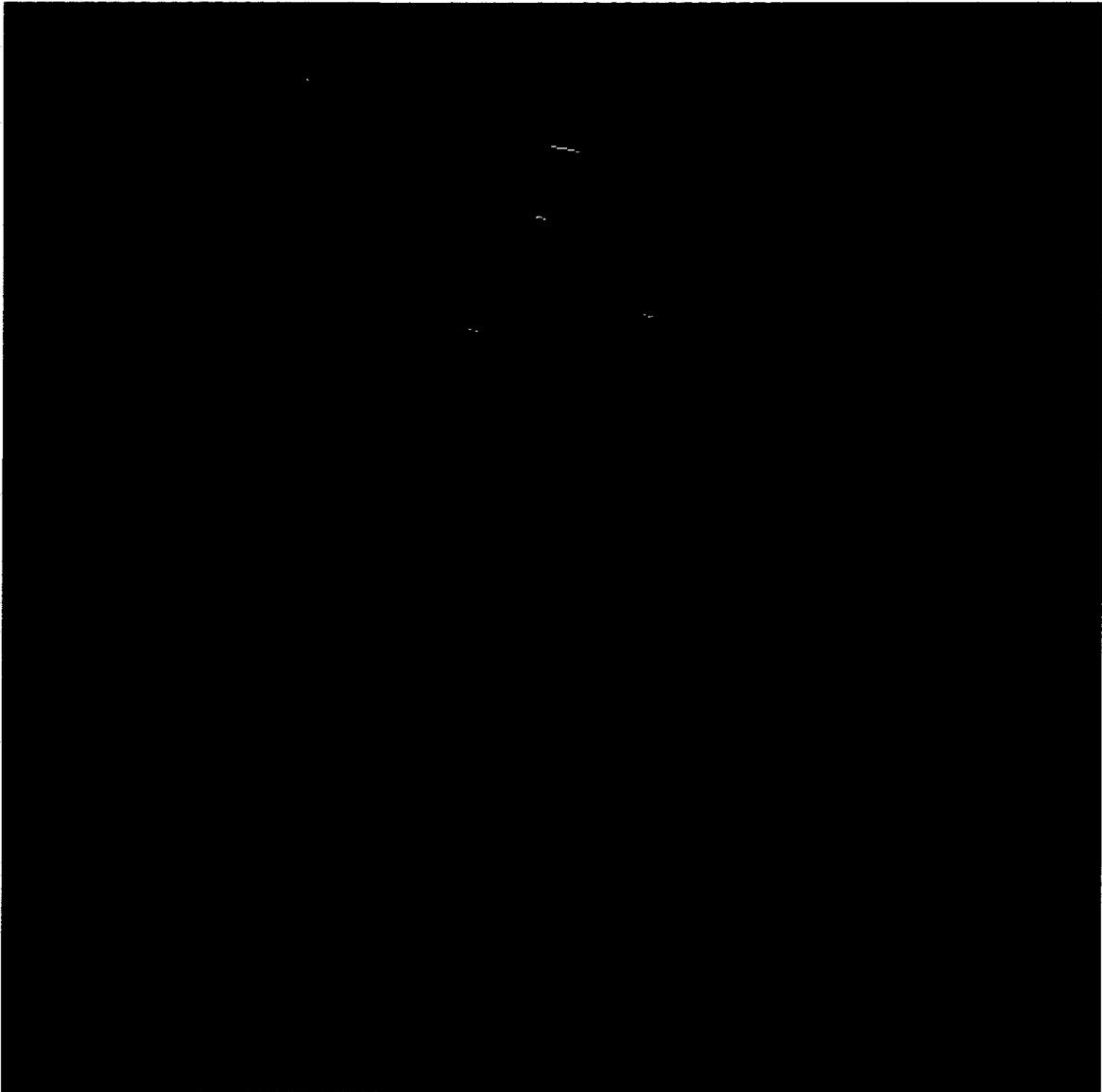


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segment.

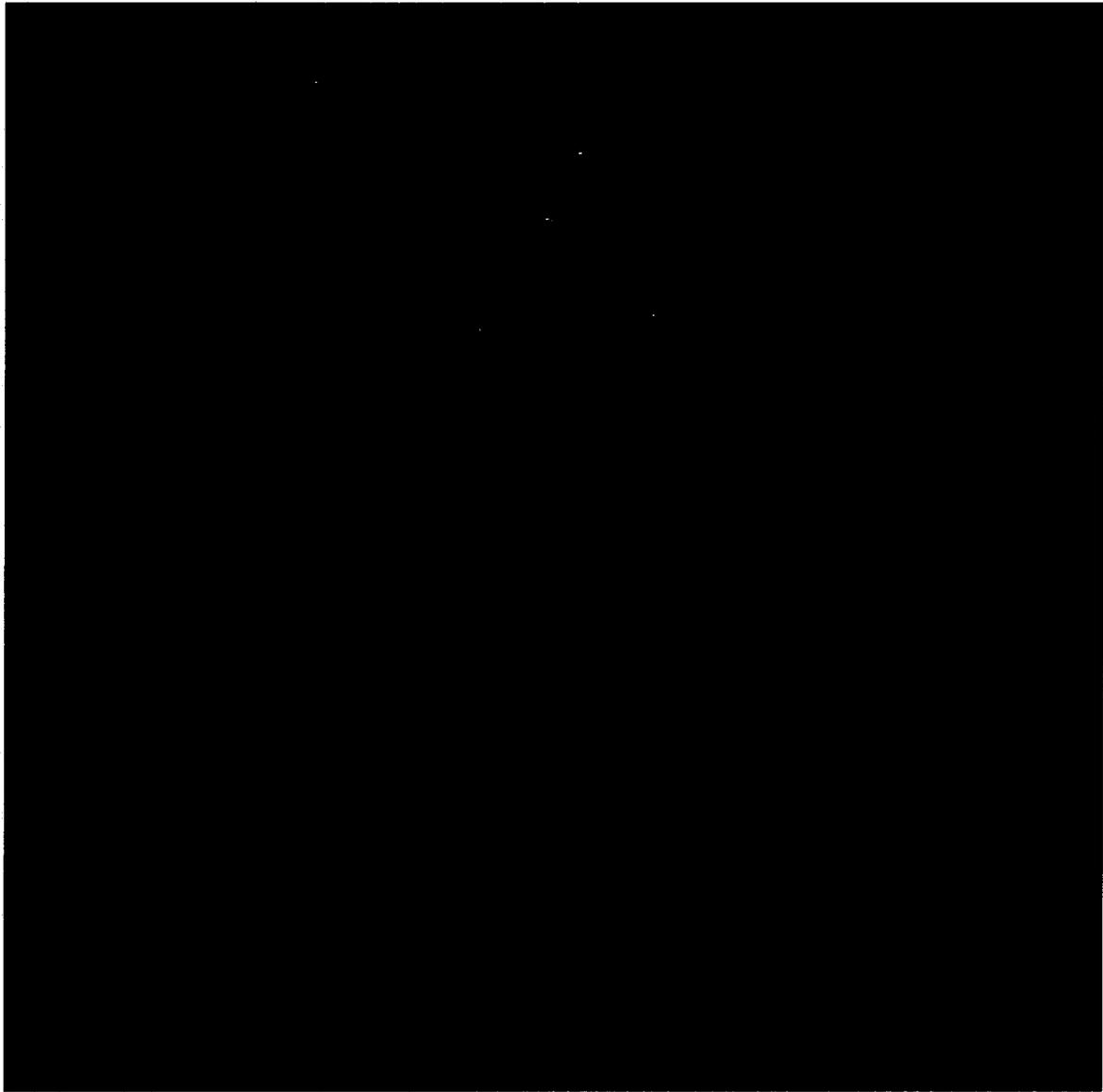


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments.

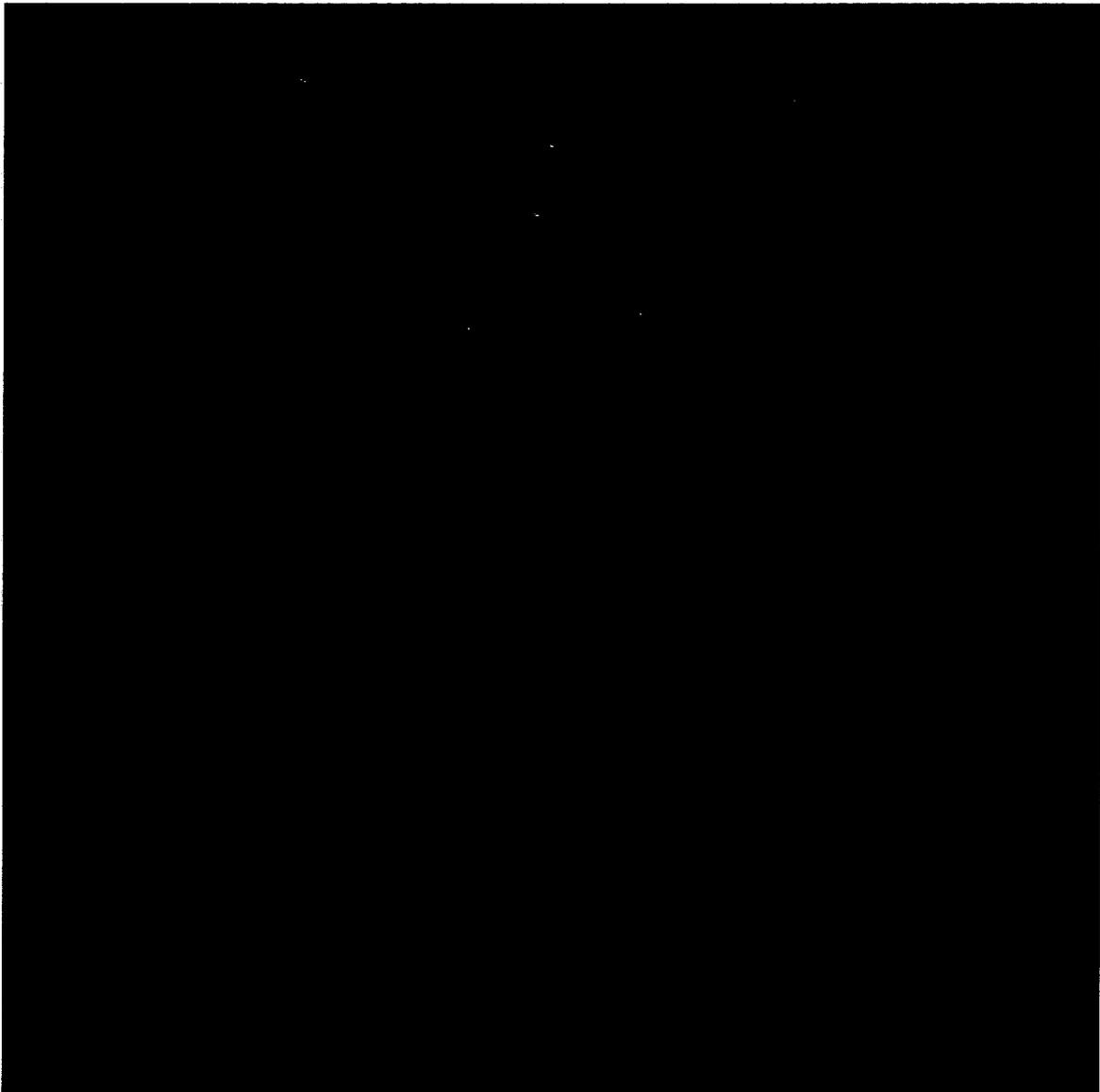
10.4.5. -15 Degree Results



Small line segments aligned in the matching orientation are correctly identified.

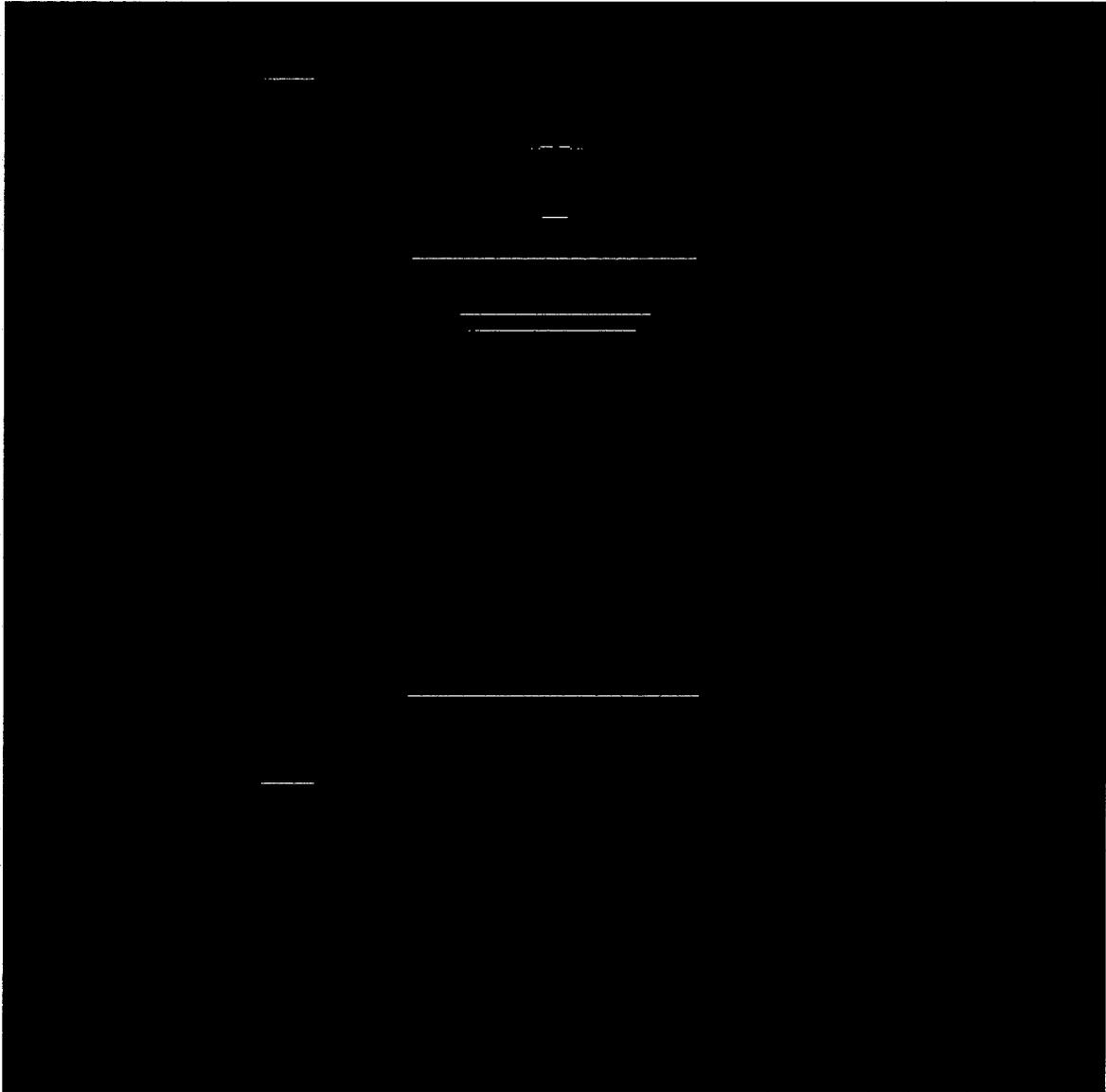


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments.

10.4.6. 0 Degree Results



Small line segments as well as long edges which are aligned in the matching orientation are correctly identified.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edges.

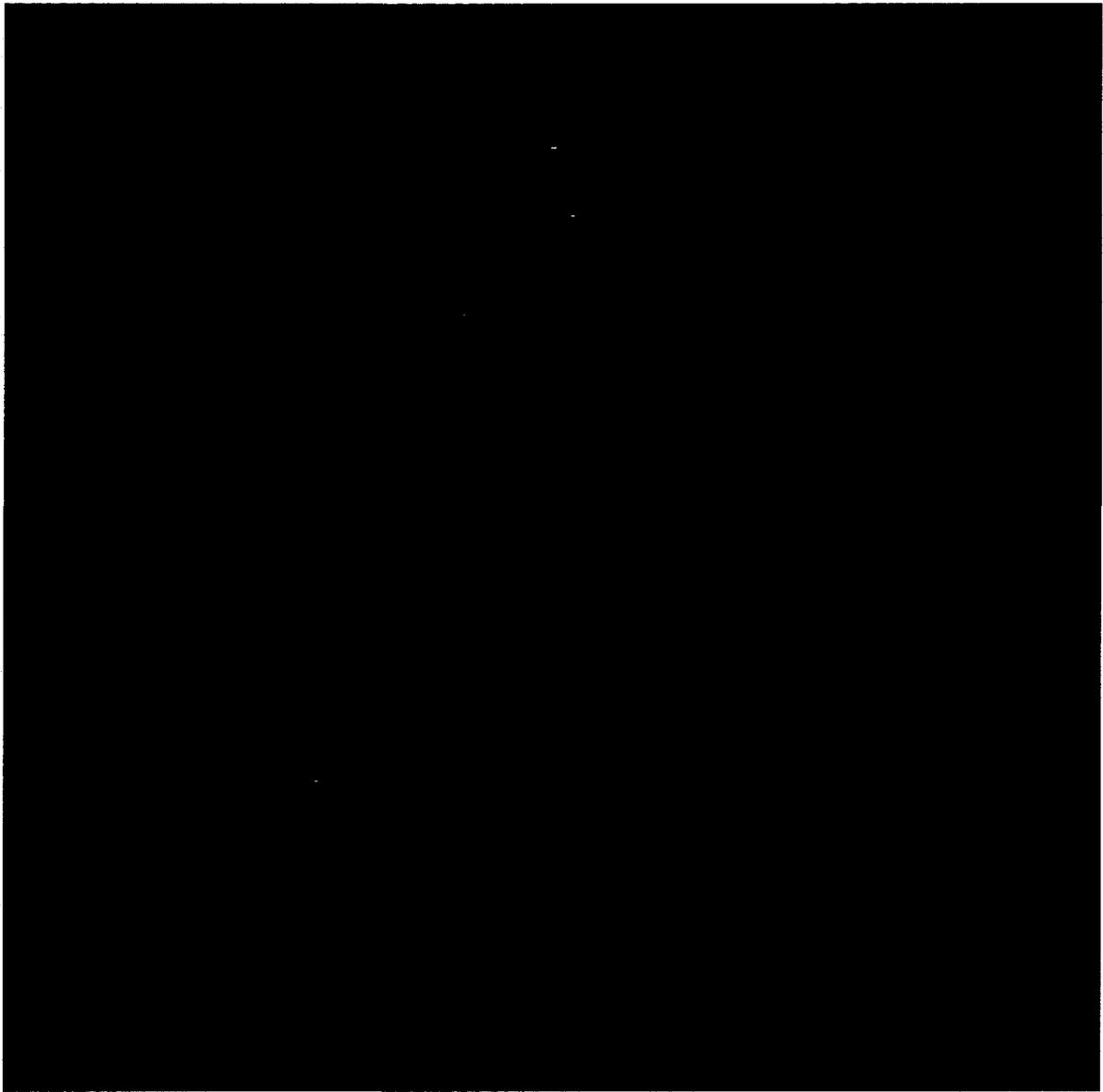


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edges.

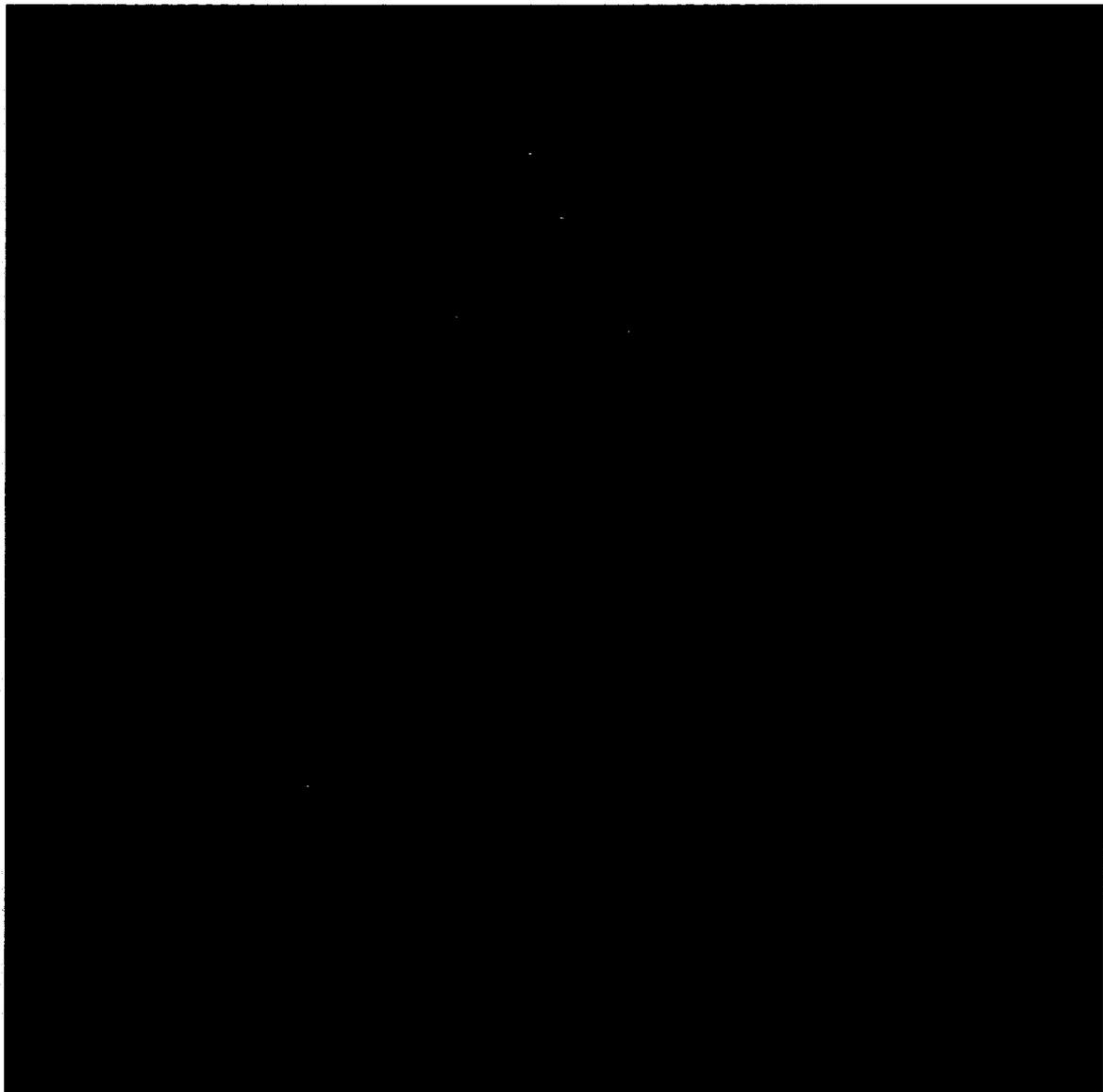
10.4.7. 15 Degree Results



Small line segments aligned in the matching orientation are correctly identified. As well, as small amount of stray activation occurs.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments. The stray activation of the Complex Cell is not sufficient to produce activation of the End Stopped Cells, and cannot affect further processing.

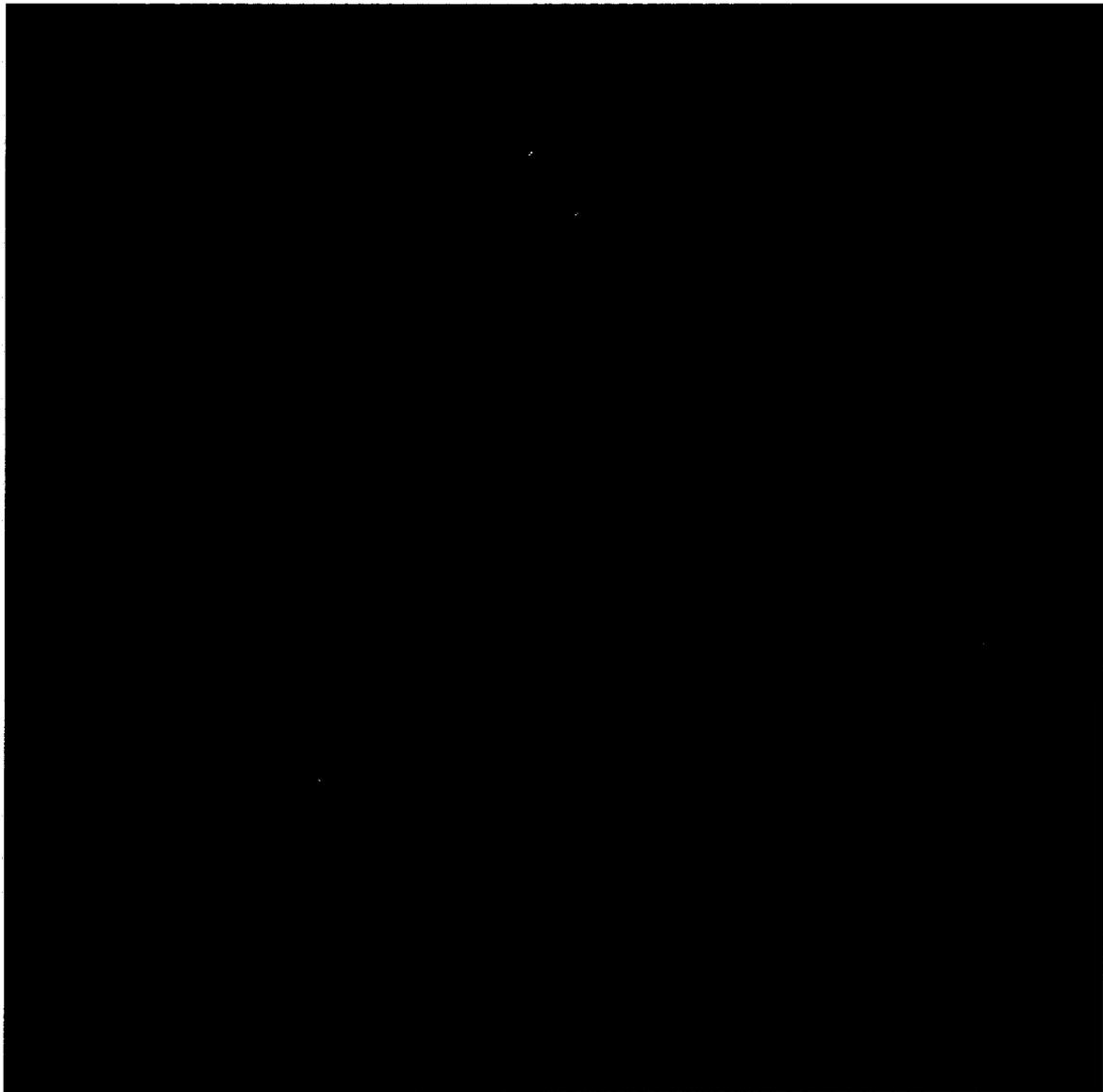


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments. The stray activation of the Complex Cell is not sufficient to produce activation of the End Stopped cell, and cannot affect further processing.

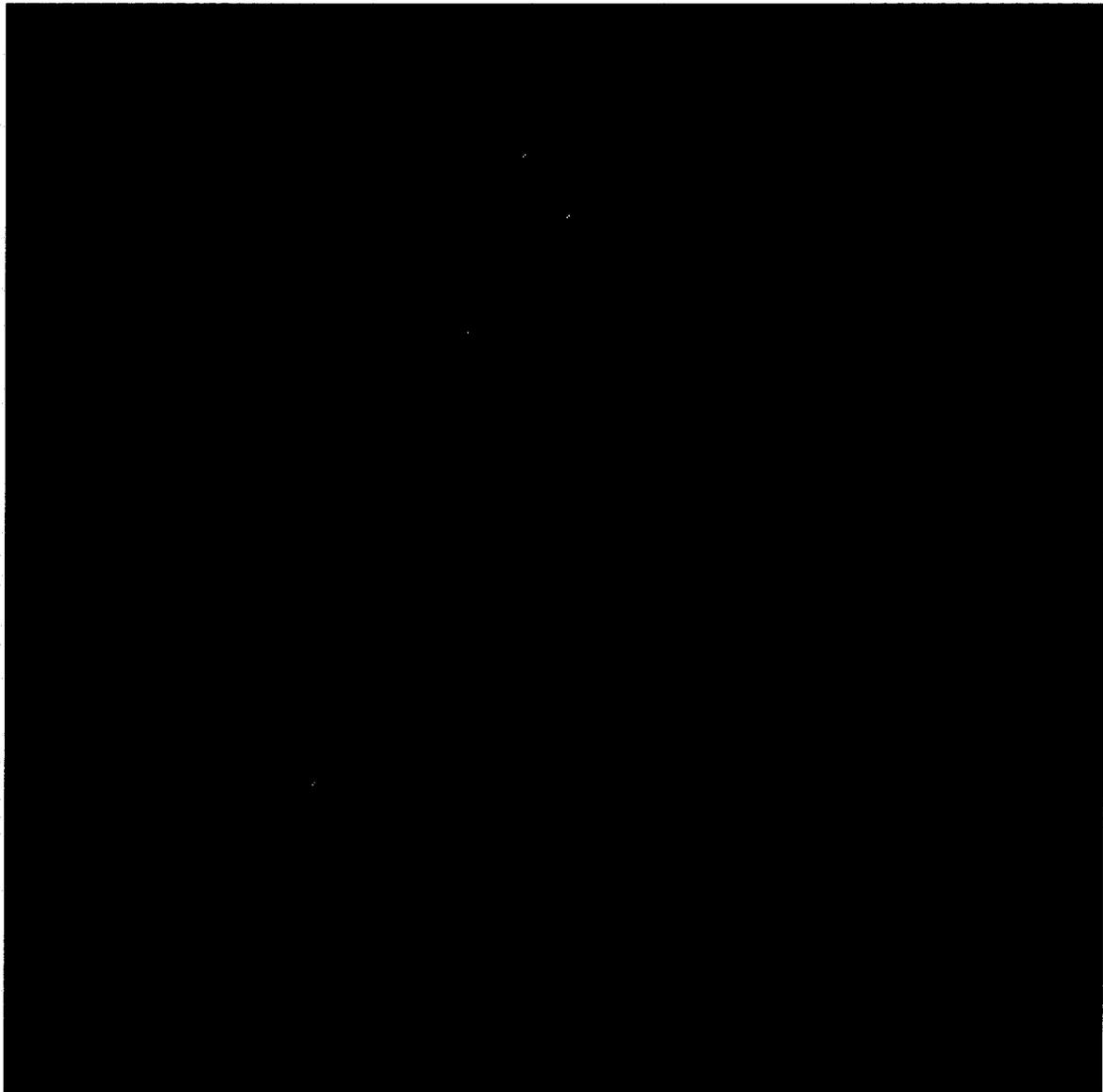
10.4.8. 30 Degree Results



Small line segments aligned in the matching orientation are correctly identified.

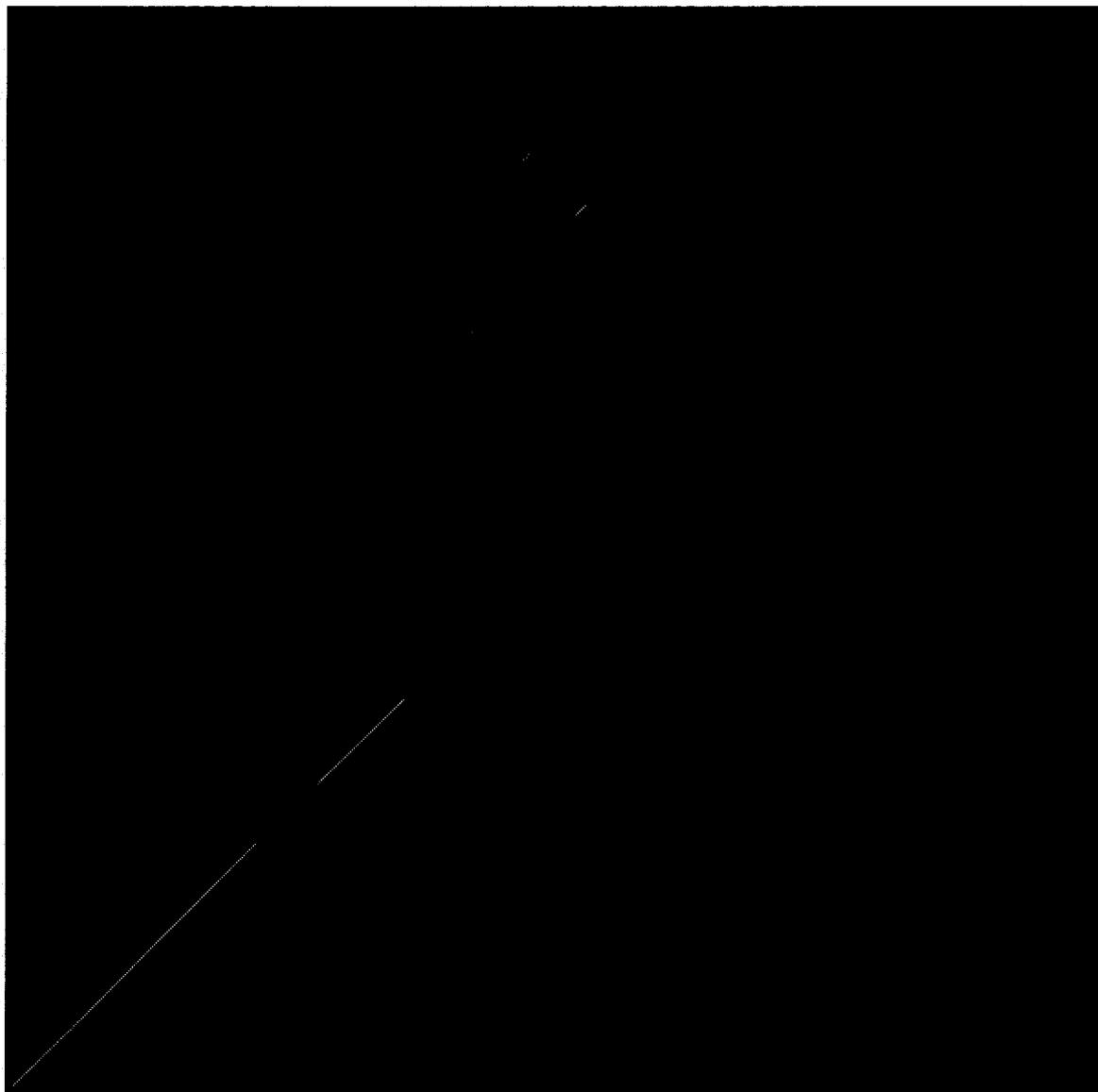


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments.

10.4.9. 45 Degree Results



Small line segments as well as edges which are aligned in the matching orientation are correctly identified.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments as well as the edges.

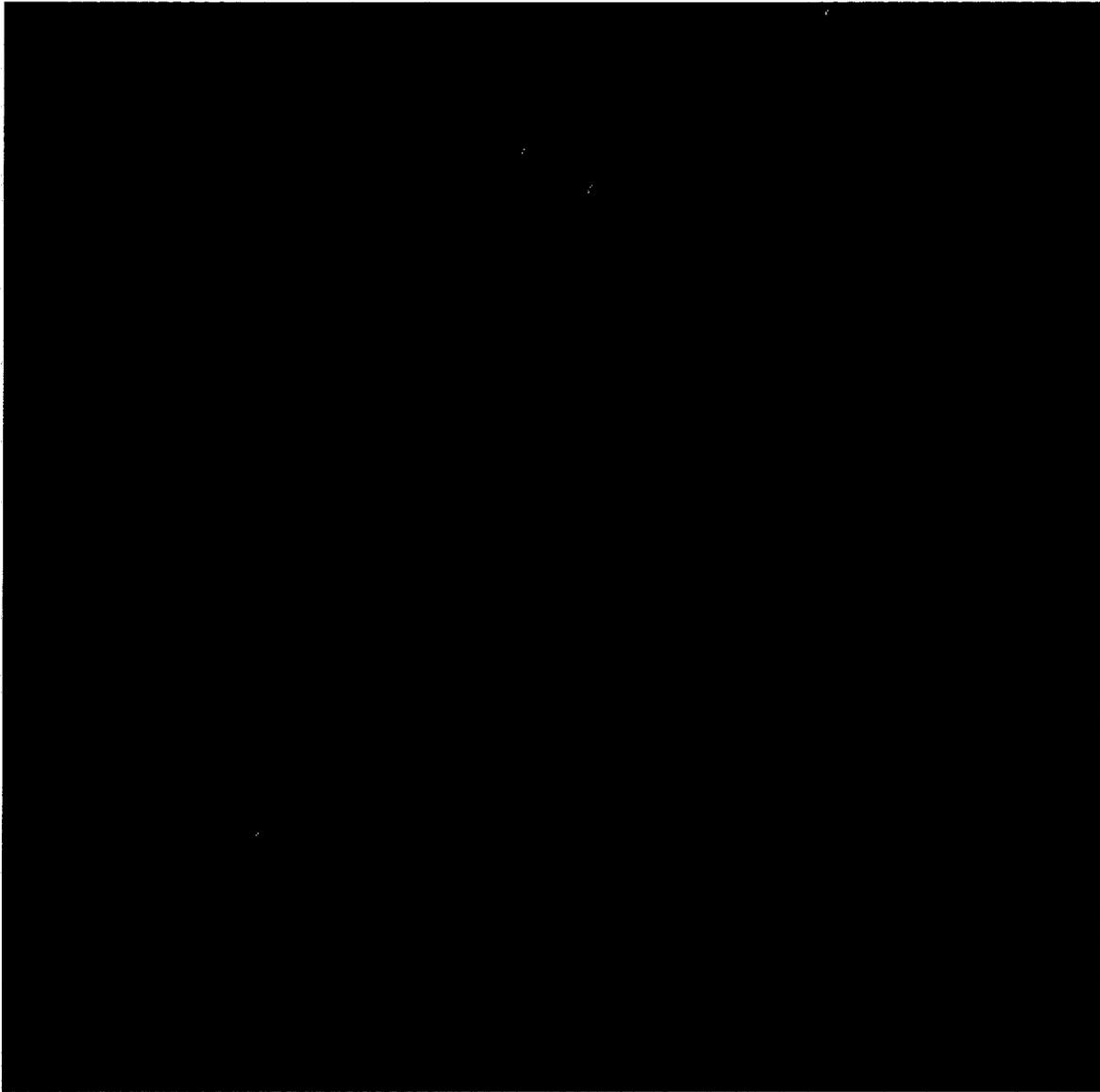


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments as well as the edges.

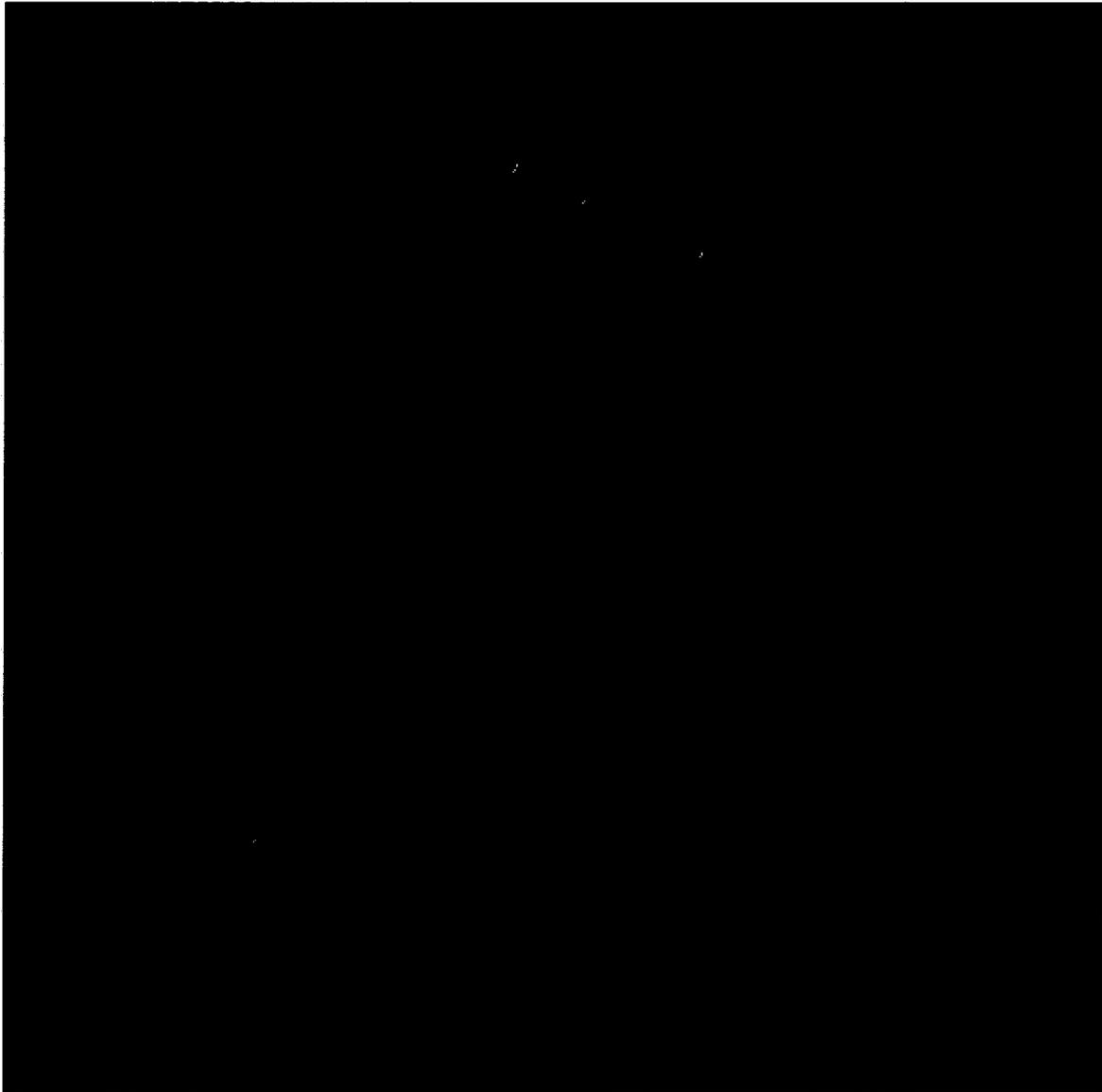
10.4.10. 60 Degree Results



Small line segments as well as long edges which are aligned in the matching orientation are correctly identified.

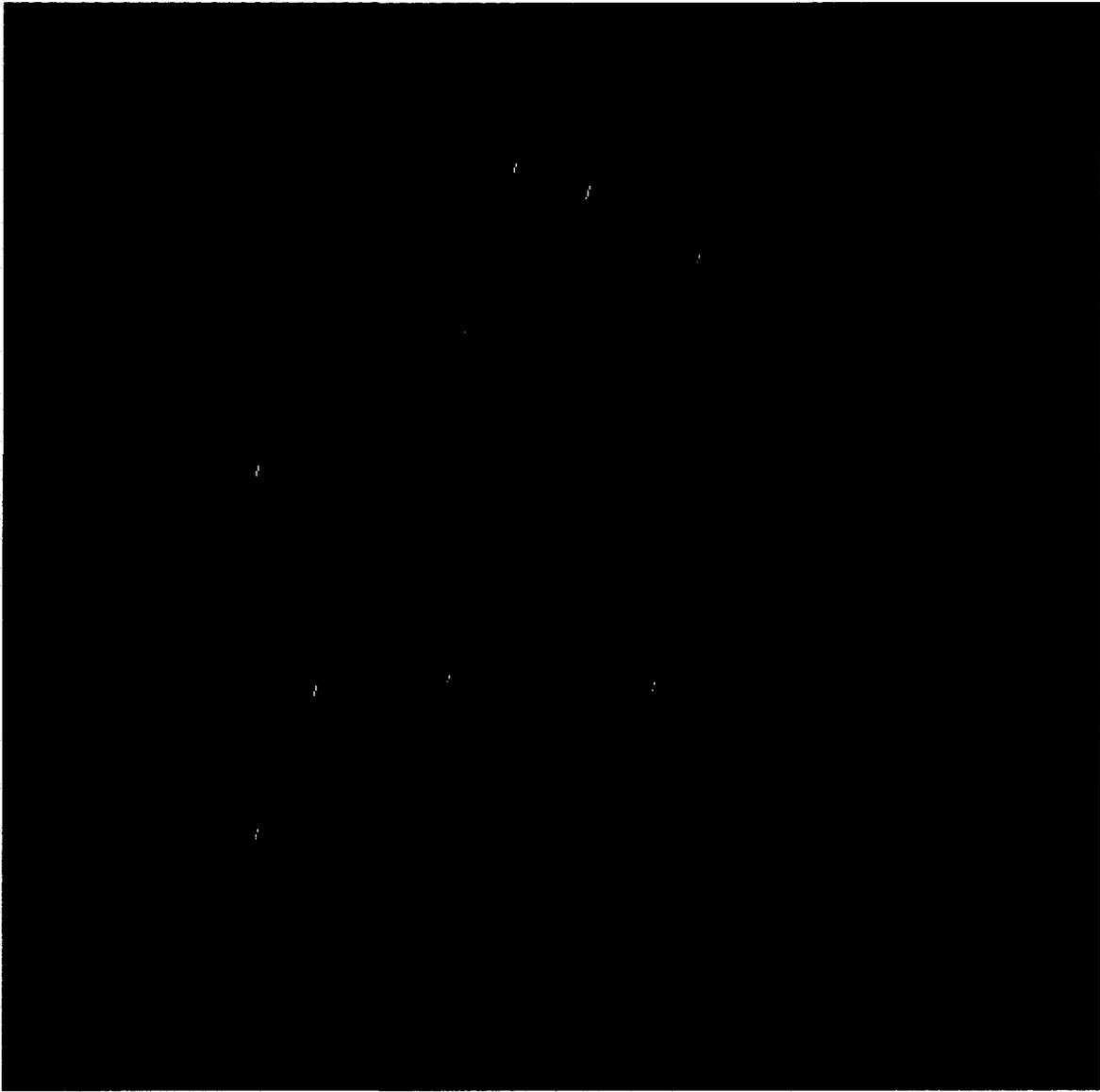


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edges.

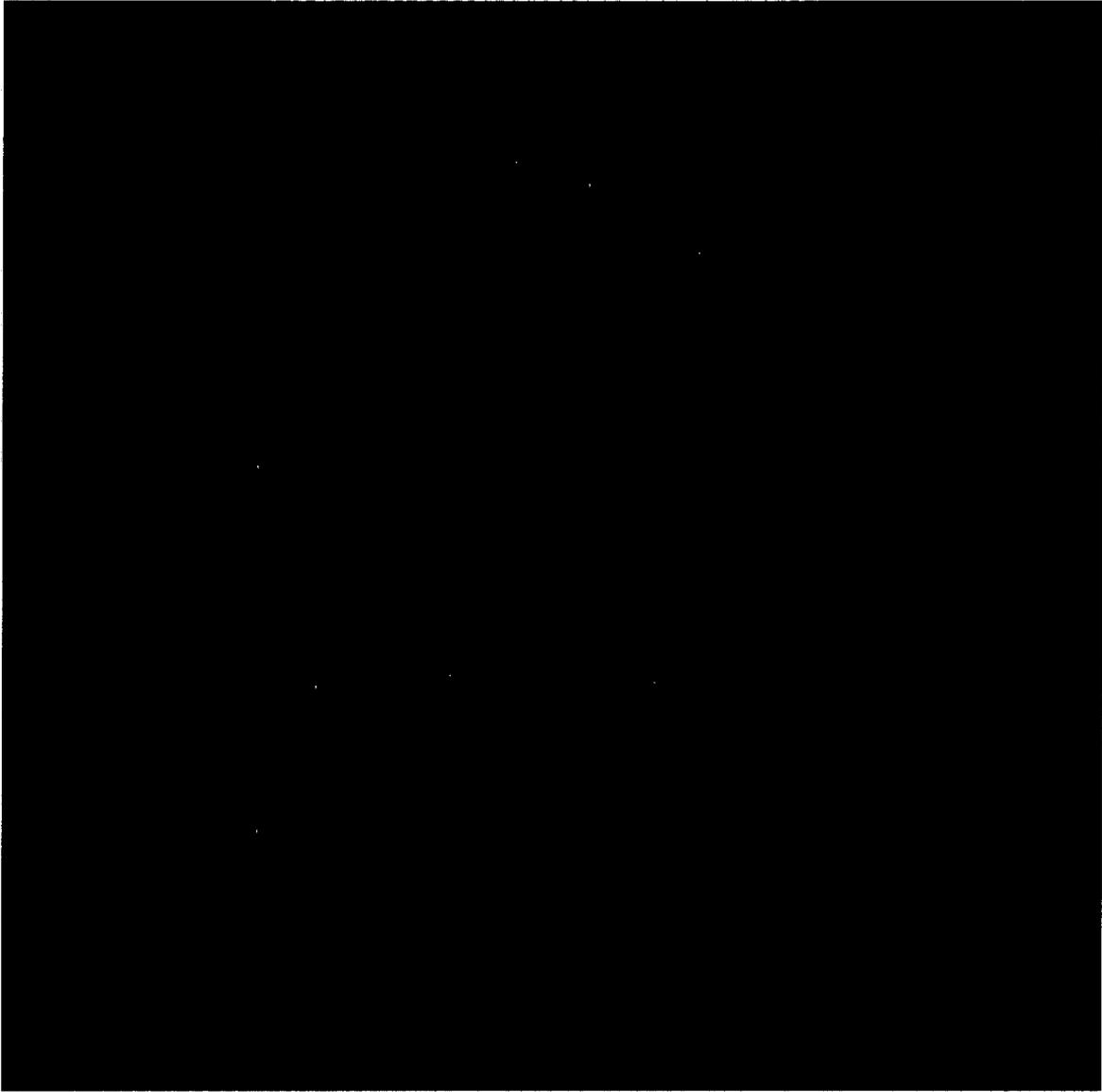


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edge.

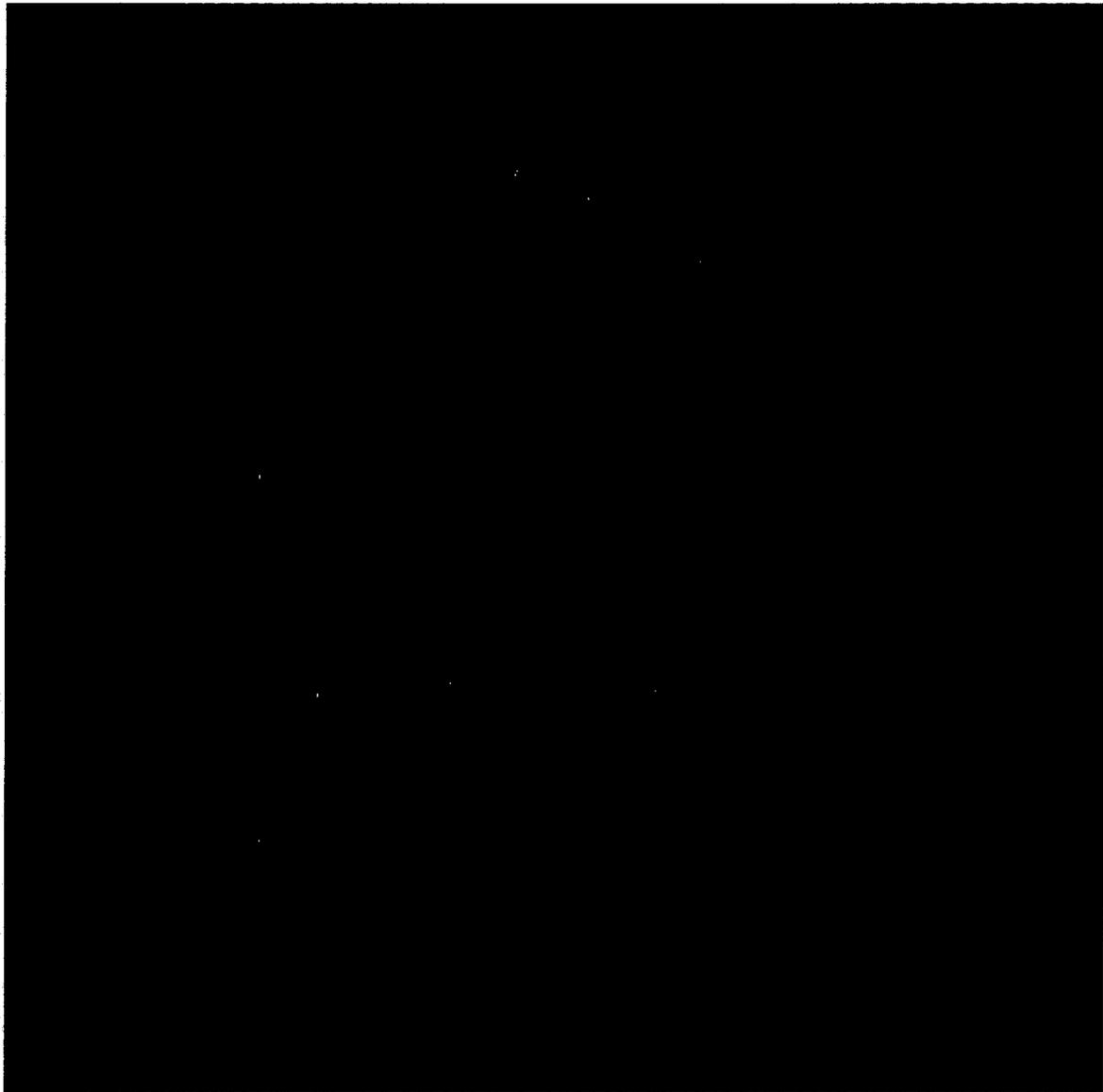
10.4.11. 75 Degree Results



Small line segments aligned in the matching orientation are correctly identified. In addition, an edge with orientation close to the matching cell's orientation is partially identified.

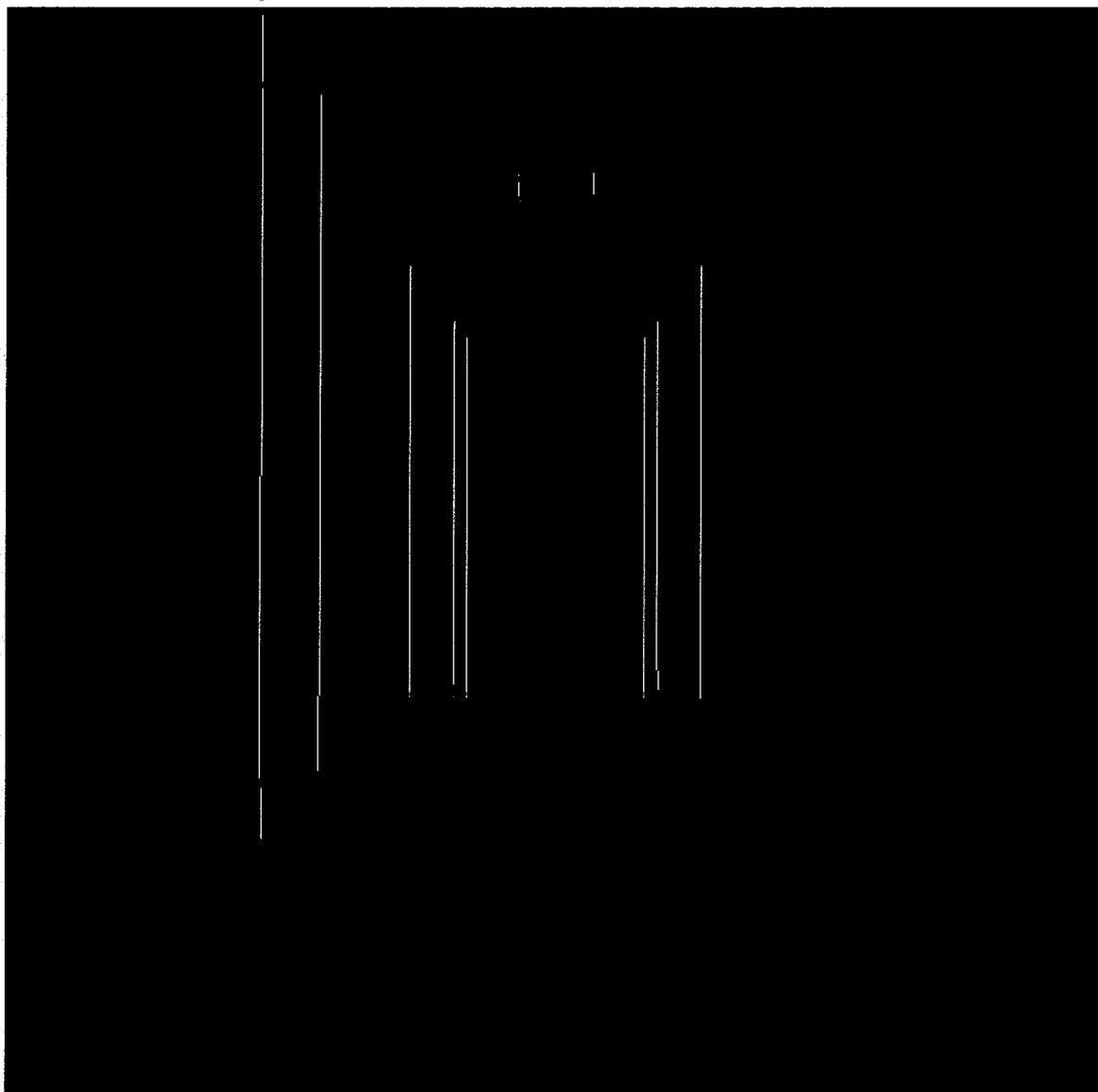


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments. The partial activation of the Complex Cell is not sufficient to produce activation of the End Stopped Cell.

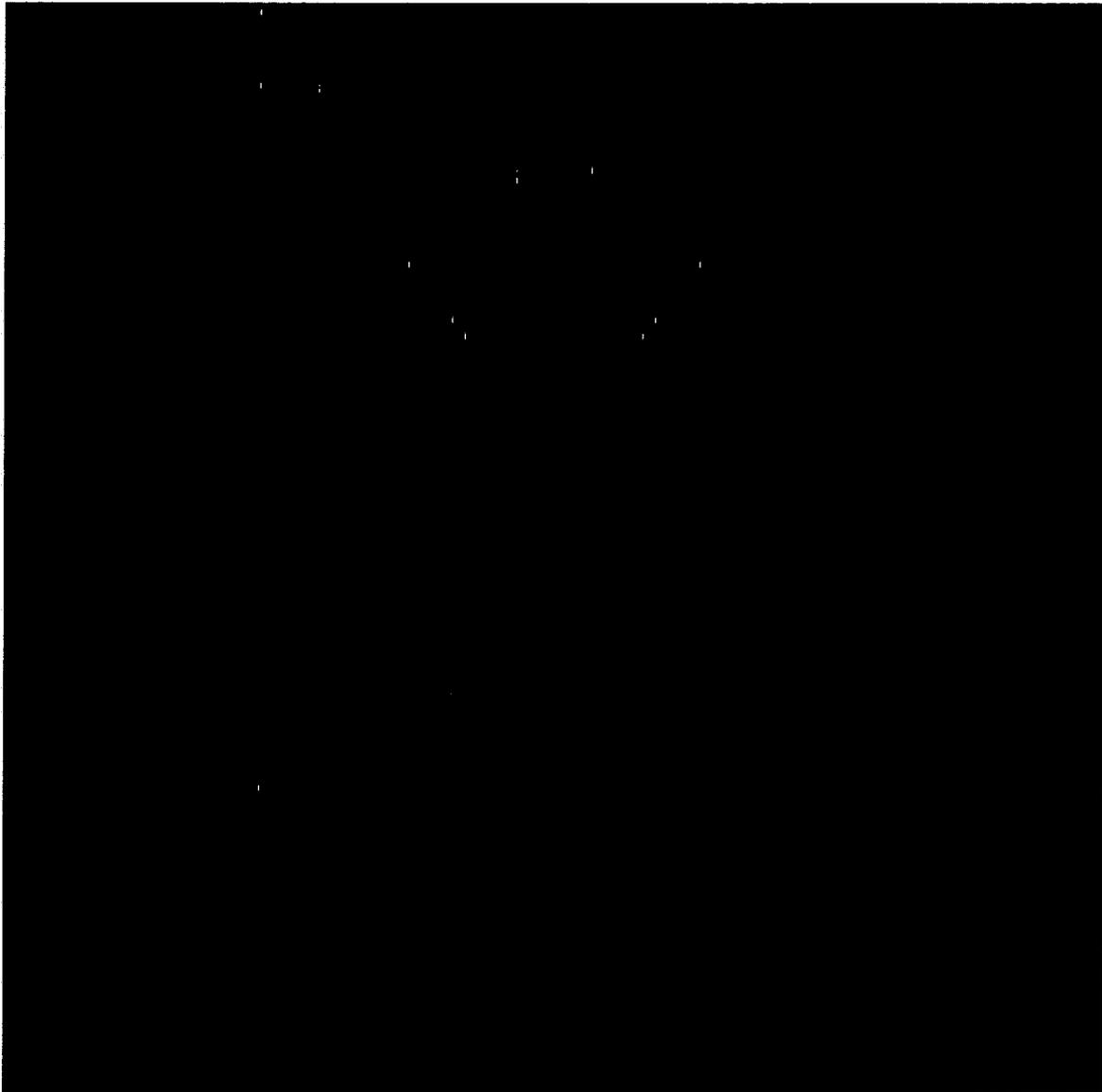


Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments. The partial activation of the Complex Cell is not sufficient to produce activation of the End Stopped Cell.

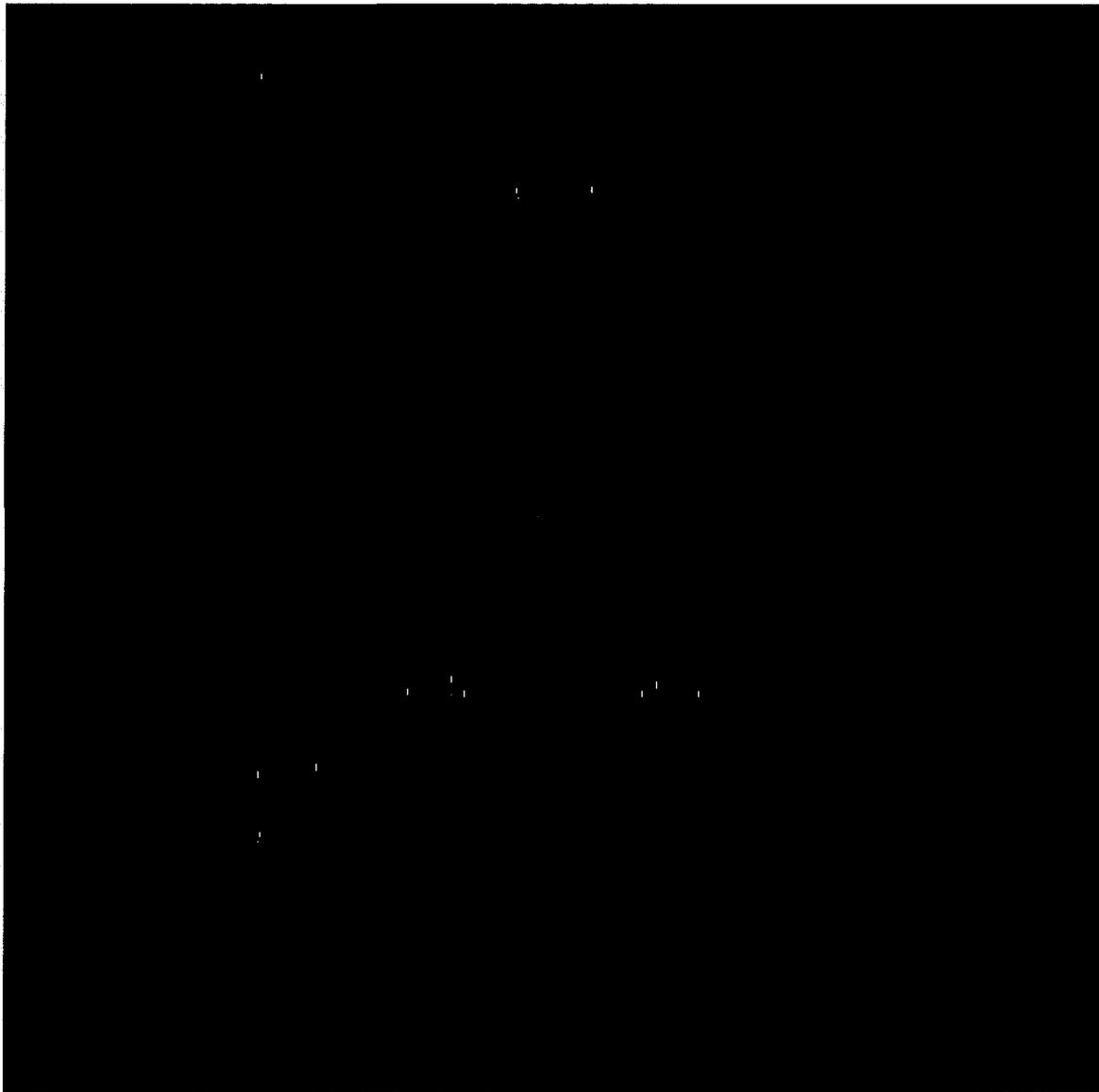
10.4.12. 90 Degree Results



Small line segments as well as long edges which are aligned in the matching orientation are correctly identified.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edges.



Similarly oriented End Stopped Cells correctly identify the start and end of the identified line segments and edges.

11. References

11.1. Papers

- [1] G. Adorni, G. Destri, M. Mordonini, F. Zanichelli, **Robot Self-Localization by Means of Vision**, Proceedings of the 1st Euromicro Workshop on Advanced Mobile Robots, pp 160-165, © 1996 IEEE
- [2] Giovanni Adorni, Stefano Cagnoni, Monica Mordonini, **Landmark-Based Robot Self-Localization: A Case Study for the RoboCup Goal-Keeper**, 1999 International Conference on Information Intelligence and Systems, © 1999 IEEE
- [3] Sven Behnke, Raun Rojas, **Neural Abstraction Pyramid: A Hierarchical Image Understanding Architecture**, IEEE International Conference on Neural Networks, Vol 2, 1998, pp 820 – 825
- [4] Sven Behnke, **Hebbian Learning and Competition in the Neural Abstraction Pyramid**, International Joint Conference on Neural Networks, 1999, pp 1356-1361
- [5] Sven Behnke, **Learning Iterative Image Reconstruction**, International Joint Conference on Artificial Intelligence, 2001, pp 1354-1358
- [6] Virginio Cantoni, Alfredo Petrosino, **2-D Object Recognition by Structured Neural Networks in a Pyramidal Architecture**, Proceedings of the 5th IEEE International Workshop on Computer Architectures for Machine Perception, © 2000 IEEE

- [7] Antonella Carbonaro, Primo Zingaretti, **Landmark Matching in a Varying Environment**, Proceedings of the Second Euromicro Workshop on Advanced Mobile Robots, pp 147-153, © 1997 IEEE
- [8] Cyril Caucois, Eric Brassart, Bruno Marhic, Cyril Drocourt, **An Absolute Localization Method Using a Synthetic Panoramic Image Base**, Proceedings of the Third Workshop on Omnidirectional Vision, © 2002 IEEE
- [9] Ee-Chien Chang, Chee K Yap, **A Wavelet Approach to Foveating Images**, Computational Geometry 97, pp 397 – 399, © 1997 ACM
- [10] Deiter Fox, Wolfram Burgard, Sebastian Thrun, **Markov Localization for Mobile Robots in Dynamic Environments**, Journal of Artificial Intelligence Research 11 (1999), pp 391-427, © 1999 AI Access Foundation and Morgan Kaufman Publishers
- [11] Jeurgen Gausemeier, Beat Bruederlin, **Development of a Real Time Image Based Object Recognition Method for Mobile AR-Devices**, Proceedings of the 2nd International Conference on Computer Graphics, Virtual Reality, Visualization and Interaction in Africa, pp 133-139, © ACM
- [12] Alexa Hauk, Stefan Kabser, Christoph Zierl, **Hierarchical Recognition of Articulated Objects from Single Perspective Views**, Proceedings on the 1997 Conference on Computer Vision and Pattern Recognition, pp 870-876, © 1997 IEEE

[13] Ramana Isukapalli, Russell Greiner, **Efficient Interpretation Policies**, International Joint Conference on Artificial Intelligence, 2001, pp 1381-1387

[14] Jason A Janet, Ricardo Gutierrez-Osuna, Troy A Chase, Mark White, Ren C Luo, **Global Self-Localization for Autonomous Mobile Robots Using Self-Organizing Kohonen Neural Networks**, Proceedings of the International Conference on Intelligent Robots and Systems (1995), pp 504-509, © 1995 IEEE

[15] Todd M Jochem, Dean A Pomerleau, Charles E Thorpe, **Vision-Based Neural Network Road and Intersection Detection and Traversal**, Proceedings of the International Conference on Intelligent Robots and Systems, pp 344-349, © 1995 IEEE

[16] Eric Marchand, Francois Chaumette, **Controlled Camera Motions for Scene Reconstruction and Exploration**, Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition, pp 169-176, © 1996 IEEE

[17] Eric Marchand, Francois Chaumette, **Active Vision for Complete Scene Reconstruction and Exploration**, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 21, No 1, January 1999, pp 65-72, © 1999 IEEE

[18] Jose del R. Millan, Angelo Arleo, **Neural Network Learning of Variable Grid Based Maps for the Autonomous Navigation of Robots**, Proceedings of the 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation, © 1997 IEEE

[19] Silviu Minut, Sridhar Mahadevan, **A Reinforcement Learning Model of Selective Visual Attention**, Proceedings of the 5th International Conference on Autonomous Agents, pp 457-164, © 2001 ACM

[20] Philippe Mulheim, Wee Kheng Leow, Yoong Keok Lee, **Fuzzy Conceptual Graphs for Matching Images of Natural Scenes**, International Joint Conference On Artificial Intelligence, 2001, pp 1397-1402

[21] Robert Osada, Thomas Funkhouser, Bernard Chazelle, David Dobkin, **Shape Distributions**, ACM Transactions on Graphics Vol 21, No 4, pp 807-832, © 2002 ACM

[22] Leonardo Romero, Eduardo Morales, Enrique Sucar, **An Hybrid Approach to Solve the Global Localization Problem from Indoor Mobile Robots Considering Sensor's Perceptual Limits**, International Joint Conference on Artificial Intelligence, 2001, pp 1411-1416

[23] Franc Solina, Ruzena Bajcsy, **Recovery of Parametric Models from Range Images: The Case for Superquadratics with Global Deformations**, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 21, No 2, February 1990, © 1990
IEEE

[24] Paul Suetens, Pascal Fua, Andrew J Hanson, **Computational Strategies for Object Recognition**, ACM Computing Strategies, Vol 24, No 1, March 1992, pp 5-61 © 1992
ACM

- [25] Glenn Wasson, David Kortenkamp, Eric Huber, **Integrating Active Perception with an Autonomous Robot**, Proceedings on the 2nd International Conference on Autonomous Agents, © 1998 ACM
- [26] Zhaoping Li, Joseph J. Atick, **Towards a Theory of the Striate Cortex, Neural Computation**, Vol6, Number 1, Jan 1994, pp 125-146.
- [27] Trevor Mundel, Alexander Dimitrov, Jack D Cowan, **Visual Cortex Circuitry and Orientation Tuning**, Advances in Neural Information Processing Systems, volume 9, 1997, pp 887-893 © Cambridge MA: MIT Press.
- [28] F. Heitger, L. Rosenthaler, R von der Heydt, E. Peterhans, O. Kubler, **Simulation of neural contour mechanisms: From simple to end-stopped cells**, Vision Research 32, pp 963-981, 1992.
- [29] Friedrich Heitger, Rudger von der Heydt, Esther Peterhans, Lukas Rosenthaler, Olaf Kuber, **Simulation of neural contour mechanisms: representing anomalous contours**, Image and Vision Computing 16, 1998, pp 407-421
- [30] Rolf P. Wurtz, Tino Lourens, **Corner Detection in color images by multiscale combination of end-stopped cortical cells**, ICANN '97, Lecture Notes in Computer Science vol. 1327, pp 901-906 © 1997 Springer Verlag.

[31] David G. Lowe, **Towards a Computational Model for Object Recognition in IT Cortex**, First IEEE International Workshop on Biologically Motivated Computer Vision, Seoul, Korea (May 2000).

[32] David G. Lowe, **Object Recognition from Local Scale-Invariant Features**, ICCV, 1999

[33] Edmond T. Rolls, Nicholas C. Aggelopoulos, Fashan Zheng, **The Receptive Fields of Inferior Temporal Cortex Neurons in Natural Scenes**, *The Journal of Neuroscience*, January 2003 vol. 23, pp 339-348.

[34] Alexander Dimitrov, **Spatial Decorrelation in Orientation-Selective Cortical Cells**, *Neural Computation* 10, 1998, pp 1779-1795

[35] O. Henricsson and F. Heitger, **The Role of Key-Points in Finding Contours**, in *Computer Vision -- ECCV'94*, edited by J. O. Eklundh, vol. II, pp. 371--383, Springer-Verlag, Berlin (1994).

11.2. Human Spatial and Visual Systems Textbooks

[36] Kolb, Whishaw, **Fundamentals of Human Neuropsychology 4th Ed, Chapter 19**, pp 438-464, © 1980, 1985, 1990, 1996 by W. H. Freeman and Company

[37] Stanly Coren, Lawrence M. Ward, James T. Enns, **Sensation and Perception 6th Ed**, 2004, John Wiley and Sons Inc.

11.3. Vision and Artificial Intelligence Textbooks

[37] Ra mesh Jain, Rangachar Kasturi, Burial G. Schunck, **Machine Vision**, © 1995 by McGraw-Hill, Inc

[38] George F Luger, William Stubblefield, **Artificial Intelligence 3rd Ed**, © 1998 Addison Wesley Long man Inc

[39] Milan Sonia, Vaclav Hlavac, Roger Boyle, **Image Processing, Analysis, and Machine Vision 2nd Ed**, © 1999 by Brooks/Cole Publishing Company

[40] Patrick Henry Winston, **Artificial Intelligence, 3rd Ed**, © 1992 Addison-Wesley

[41] David Marr, **Vision**, 1982, WH Freeman and Company

[42] John F. Sowa, **Conceptual Structures: Information Processing in Mind and Machines**, © 1984 Addison Wesley, Reading, MA